

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 03:30:48 ; Search time 3690 Seconds  
(without alignments)  
2981.185 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111950\_8301/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bite -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894 @CGN\_1\_1\_3437\_@runat\_15062005\_111950\_8301 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1521	100.0	960	CC344833	CG240397 OGIAC20TV
2	1406.5	92.5	923	9	CG240397 OGIAC20TV
3	1256	82.6	818	9	CG453578 OGSAD40TV
4	1248	82.1	873	8	CG344824 OGIAC20TV
5	1183	77.8	867	9	CC724960 OGIAC20TV
6	1149	75.5	719	9	CG284572 OGIAC20TV
7	1123	73.8	801	6	CA134480 SCUFRT106
8	1034	68.0	656	6	CA245555 SCBGL508
9	993.5	65.3	722	6	CA130685 SCCRT100

10	992	65.2	591	9	CG776236
11	987	64.9	671	6	CA202412 SCRLFL100
12	981	64.5	3374	3	AY109355 Zea mays
13	976	64.2	722	9	CC724950 OGIAB23TH
14	932	61.3	744	6	CA130686 SCCRT100
15	926	60.9	1191	7	CK211413 FGAS02325
16	921.5	60.6	593	8	BZ774817 1151a11.b
17	920	60.5	960	9	CG073495 PUJCG54TD
18	903	59.4	955	8	CC384766 PUHOG53TD
19	869.5	57.2	794	2	BF617713 HVSMEC001
20	811.5	53.4	907	9	CG240387 OGYC127TH
21	807	53.1	575	9	CG734025 1119161GO
22	770	50.6	619	6	CA204213 SCAGFL109
23	758	49.8	1099	7	CK215908 FGAS02788
24	727	47.8	1123	7	CK214494 FGAS02642
25	724.5	47.6	588	5	BU999708 H115H04r
26	724	47.6	910	7	CK253394 EST7737031
27	714	46.9	551	8	BH408067 1007056A0
28	705.5	46.4	1119	3	CNSOALFK
29	666.5	43.8	1007	3	CNSO926Q
30	664.5	43.7	1813	3	CNSOALNR
31	644	42.3	557	2	BF277833 GA_EB003
32	642	42.2	831	9	CG553151 MBEFW86TF
33	640.5	42.1	1431	3	CNSO92LQ
34	640	42.1	620	5	BQ620191 TAlr1176F
35	617	40.6	804	7	CNS23418 GQ015M12
36	613	40.3	453	4	BI956405 HVSME000
37	606	39.8	433	4	BM099879 E8es01 SQ
38	591	38.9	883	7	CK934794 CGF100422
39	585.5	38.5	752	7	CO865620 Mdb5012p
40	584	38.4	448	8	CG028437 3591_1_10
41	581.5	38.2	478	4	BG417083 HVSMEK001
42	580	38.1	876	9	CG073494 PUJCG54TB
43	567	37.3	409	4	BM099878 E8es01 SQ
44	562	36.9	927	9	CG284580 OGIWIF5TV
45	559	36.8	431	9	CG776505 1123005E0

#### ALIGNMENTS

RESULT 1

CC344833

LOCUS

DEFINITION

CG1AQ20TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0359C15,

genomic survey sequence.

CC344833

VERSION

CC344833.1 GI:30814239

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 960)

Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..960

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

FEATURES

source

```
/clone="ZMMBMA0359C15"  
/clone_lib="ZM_0.7_1.5_KB"  
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"
```

## ORIGIN

```
Alignment Scores:  
Pred. No.: 2,95e-150 Length: 960  
Score: 1521.00 Matches: 289  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0
```

US-10-042-894A-8 (1-289) x CC344833 (1-960)

```
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 18 ATGTCGACCTCCACCGCGGAGCACCAAGTCGCGCGCCACCGCGCTCCGCCAGCAG 77  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyThrLysProLeuGlnAlaGlyAspArg 40  
Db 78 CTGGGCGCGCTCATCGACGGCTCCGGCTCTTCTACAAGCCGCTCCAGCGCGCGACCGT 137  
QY 41 GlyGluHisGluValAlaPheTyThrGluAlaPheSerAlaHisAlaValProAlaArg 60  
Db 138 GGGAGACGAGGTGCGCTTCTATAGGCGTCTCCGCCACCGCGCGCTCCGCCCGCC 197  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 198 ATCCGAGACACCTTCTTCCCGCGGTTCACGCGACGACTCTCTCCACCGCGCGCAG 257  
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
Db 258 CCGGGGAGCGCATCTCCACCTCGTCTTCGACGACTCTCTCGCGGGTTTCAGGCGCC 317  
QY 101 CysValAlaAspIleValGlyAlaIleThrTrpProSerSerProGluProTyr 120  
Db 318 TCGTTCGACAGATCAAGATCGCGCCATCATCGTGGCCACCGAGTTCCGCGGAGCGCTAC 377  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 378 ATCCCAAGTGCTCGCCCAAGGACCGCGGACCCAGCGGCTTCTGCTCGATTCCGCGTC 437  
QY 141 SerGlyValArgValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 438 TCGGCGTCCGAGTCGTCGCGCCCGAGGCGCGCGTGTGCGGACGAGCGCGCGAGGTG 497  
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180  
Db 498 AAGGCATAGACACCGCGCGGTGCGCGCGGTCTCCGCGGTACGTGTCATCCGTTGCC 557  
QY 181 AspGluGlyMetAspCysAlaLeuAlaValTyrGlyLysGlyGlyValLeu 200  
Db 558 GACGAGGGATGAGCTGTGGCTCGCGCGCGGTGTACGGAGAAAGTGGAGTCTTG 617  
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluThrLeuPheHisPheTyThr 220  
Db 618 TCACAGCTCGCGAGCTCAAGCGGTGTTTCGAGGACGAGACTGTGTTCCACTTCTACTCG 677  
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyGlyGly 240  
Db 678 GGTCTGATTCTTCTGGGCTATGATGCTGTCAGTGCAGAGCGGAGATGGGGTGGG 737  
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260  
Db 738 GTGACGCTGAAGCTGTGTGACTTTGCCCATGTGGCCGAGGGTGTGGGTGATTGACAC 797  
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280  
Db 798 AACTTCTCGGGGGGCTCTGCTCGTGTATCAAGTTCTGTTTCTGACATTGTTCCGGAGCT 857  
QY 281 ProHisThrGlnProLeuGlyProSer 289
```

Db 858 CCTCATACGACGCTTTGGTCTCT 884

## RESULT 2

CG240397

## LOCUS

DEFINITION

CG240397 ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA078BE05,

genomic survey sequence.

ACCESSION

CG240397 GI:34140283

KEYWORDS

GSS.

SOURCE

Zea mays

## ORGANISM

Zea mays; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 923)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Reanick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGYCI27TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

## FEATURES

source

```
1..923  
/organism="Zea mays"  
/mol_type="genomic DNA"  
/strain="B73"  
/db_xref="taxon:4577"  
/clone="ZMMBMA078BE05"  
/clone_lib="ZM_0.7_1.5_KB"  
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"
```

## ORIGIN

## Alignment Scores:

Pred. No.: 3,73e-138 Length: 923

Score: 1406.50 Matches: 277

Percent Similarity: 96.21% Conservative: 2

Best Local Similarity: 95.52% Mismatches: 10

Query Match: 92.47% Indels: 2

DB: 9 Gaps: 0

US-10-042-894A-8 (1-289) x CG240397 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20

Db 34 ATGCCCGACCTCCACCGCGGAGCACCAAGTCGCGGTCCCGCGCTCCGCCAGCAG 93

QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyThrLysProLeuGlnAlaGlyAspArg 40

Db 94 CTGGCGCGCTCATCGACGGCTCCGCTCTTCTACAAGCGCTCCAGCGCGCGACCGT 153

QY 41 GlyGluHisGluValAlaPheTyThrGluAlaPheSerAlaHisAlaValProAlaArg 60

Db 154 GGGGAGACGAGGTGCGCTTCTATAGCGGTCTCCGCCACCGCGCGCTCCGCCCGCC 213

QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80

Db 214 ATCCGAGACACCTTCTTCCCGGTTCCACGCGACGCGACTCTCCACCGAGGCGCAG 273

QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100

Db 274 CCGGGGAGCGCGCATCCCGACCTCTGTCGACGACCTCTCTCGCGGGTTTGAGGCGCC 333

QY 101 CysValAlaAspIleLysAlaGlyAlaIleThrTrpProSerSerProGluProTyr 120  
Db 334 TGGTTCGAGACATCAAGATCGCGCCATCAGTGGCCACCGAGTTCCGGAGCCCTAC 393  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 394 ATCGCCAGTACCTCGCCCAAGACCGCGGACCCAGCGGTTCGTCTCGATTCGCGTC 453  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 454 TTGC--GTCCGAGTCTCGCGCCCGAGGCGCGTGTGGCGGACGCGCGGAGGTG 511  
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180  
Db 512 AAGGCTATGACACCGCTCGCGGTCCCGCGGTCTCGCGGTCTCGGTATCCGCTTCG 571  
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyLysGlyGlyValLe 200  
Db 572 CGACGAGGGGATGAGTCTCGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTT 631  
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
Db 632 GTACAGCTGCGCGAGCTCAAGGCATGGTGGAGGAGACACTCTGTTCACCTTCTACTC 691  
QY 220 xAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
Db 692 GGGCTCGATTCTTCTGGGCTATGATCTGTCGAGTCCGACAGCGAGGTGGGGGTGG 751  
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
Db 752 GGTACAGTGAAGTGGTGGACTTTGCCCATGTGGCGGAGGTGATGGGTGATTGACCA 811  
QY 260 sAenPheLeuGlyGlyLeuCysSerLeuLysPheValSerAspIleValProGluTh 280  
Db 812 CAACCTTCCTGGCGAGCTCTGTAGCTGATCAAGTTCGTTCTGACATTGTTCCAGAGAC 871  
QY 280 rProHisThrGlnProLeuGlyProSer 289  
Db 872 TCCTTAGACGACGCCCTTGGTCTCTCT 899

RESULT 3  
CG453578 818 bp DNA linear GSS 17-SBP-2003  
LOCUS CG9AD40TV ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBma0849H08,  
DEFINITION Genomic survey sequence.

ACCESSION CG453578  
VERSION CG453578.1 GI:34838578  
KEYWORDS GSS.

## SOURCE

Zea mays

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 818)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OG9AD40TH

Contact: Cathy Whitelaw

## TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..818

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

## FEATURES

## Source

CC344824 873 bp DNA linear GSS 16-MAY-2003  
LOCUS OGIAQ20TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBma0359C15,  
DEFINITION Genomic survey sequence.  
ACCESSION CC344824  
VERSION CC344824.1 GI:30814230  
KEYWORDS GSS.

/db\_xref="taxon:4577"  
/clone="ZMMBma0849H08"  
/clone\_lib="ZM 0.7\_1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,74e-122 Length: 818  
Score: 1256.00 Matches: 239  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.58% Indels: 0  
DB: 9 Gaps: 0

US-10-042-894A-8 (1-289) x CG453578 (1-818)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20

Db 96 ATGTCCGACCTCCACCCGCGGAGCACCAGTCCGCGGCACCGCGCTCCGCGCAG 155

QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40

Db 156 CTGGGCCCGCTCATCGACGGCTCCGGCCCTCTTCTACAAGCCGCTCCAGGCCGCGGACCG 215

QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60

Db 216 GGGGAGACAGAGTTCGCTTCTATGAGGGTCTTCGCCCCACCGCGCTCCGCGCGC 275

QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80

Db 276 ATCCGAGACACCTTCTTCCCGGTTCCACGGCAGCAGCTCTCTCCACCGAGGGCGAG 335

QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100

Db 336 CCGCGGAGCCGACCTCCTCACCTCGCTTCGAGACCTCTCTCGGGGTTCAGGGCGGCC 395

QY 101 CysValAlaAspIleLysAlaIleThrTrpProSerSerProGluProTyr 120

Db 396 TCGGTCCGAGACATCAAGATCGCGGCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 455

QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140

Db 456 ATCGCCAGTGGCTCGCCAAAGACCGCGGACCCAGCGGCTCTGTCGATTCGCGGTC 515

QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160

Db 516 TCGCGCTCCGAGTCTGCGGCCCGAGGGCGCGTGTGGCGGACCGAGCCCGGAGGTG 575

QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSerValAla 180

Db 576 AAGGCATCGACACCGCGCGGTCTCGCGCGTGTCTCGCGGCTACGTGTATCCGTTGCC 635

QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyLysGlyGlyValLeu 200

Db 636 GACGAGGGGATGAGTGTGCGCTCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTG 695

QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer 220

Db 696 TCACAGCTCGCGAGCTCAAGCGTGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCG 755

QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 239

Db 756 GCGTCGATTCTTCTGGGCTATGATGCTGTGTCAGTTCGACAGCGCGGAGATGGGGG 812

## RESULT 4

## CC344824/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

```
SOURCE      Zea mays
ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 873)
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
              Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
              Consortium for Maize Genomics
              Unpublished (2002)
TITLE        TIGR
JOURNAL       9712 Medical Center Drive, Rockville, MD 20850, USA
COMMENT      Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TR
              Class: sheared ends.
              Location/Qualifiers
                source
                  1..873
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone="ZMMBma0359c15"
                    /clone_lib="ZM 0.7 1.5 KB"
                    /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                    methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.:      2,11e-121      Length:      873
Score:          1248.00      Matches:    238
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    82.05%      Indels:     0
DB:             8      Gaps:         0

US-10-042-894A-8 (1-289) x CC344824 (1-873)

QY    52  SerAlaHisAlaAlaValProAlaArgIleArgAspThrPhePheProArgPheHisGly  71
      |||||
DB    872  TCGGCCACCGCGCGCTCCGGCCCGCATCCGAGACACCTTCTCCCGCGGTTCACGCG  813

QY    72  ThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAsp  91
      |||||
DB    812  ACGCGACTCTCCCAACCGAGCGCAGCCGGGGAGCGCATCTCACCTCGTCTCTCGAC  753

QY    92  AspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIleValIleGlyAlaIleThr  111
      |||||
DB    752  GACCTCTCGCGGGGTTTCAGCGCCCTCGTGCAGACATCAAGATCGCGGCCATCAG  693

QY    112  TtpProProSerSerProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThr  131
      |||||
DB    692  TGGCCACCGAGTTCGGCGGAGCCCTACATGCCAAGTGCTGCCAAGACCGCGGGACC  633

QY    132  ThrSerValLeuLeuGlyPheArgValSerGlyValArgValValGlyProGluGlyAla  151
      |||||
DB    632  ACGAGGTTCTGCTCGGATTCGCGCTCTCCGGCGTCCGAGTCTCGGCCCGCGAGCGCC  573

QY    152  ValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgVal  171
      |||||
DB    572  GTGTGGCGGACGAGCGCCCGAGGTGAAGGCCATGGACACCGCGCGCTCCGCGCGTG  513

QY    172  LeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla  191
      |||||
DB    512  CTCGGCGGTACGTGTCACTCCGTTCGCGACGAGGGATGGATGTGCGCTCCGCGCGCG  453

QY    192  ValTyrGlyGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu  211
      |||||
DB    452  GTGTACGAGGAAAGGTGGAGTCTTGTACAGCTCGCGGAGCTCAAGCGTGGTTCGAG  393

QY    212  GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAla  231
      |||||
DB    332  GAGCAGACTCTGTTCCACTTCTACTTCGGCTCGATCTTCTGGCTATGATGCTGCTGCA  333

QY    232  ValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisVal  251
      |||||
DB    332  GTCGACGACGAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTG  273

QY    252  AlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLys  271
      |||||
DB    272  GCCGAGGTGATGGGGTGATTGACCACCACTTCTTCGGCGGGCTCTGCTCGCTGATCAAG  213

QY    272  PheValSerAspIleValProGluThrProHisThrGlnProLeuGlyProSer  289
      |||||
DB    212  TTCGTTTTCGACATGTTCCGGAGACTCTCTATACGACGCTTGGGTCCTTCT  159

RESULT 5
CC724960
LOCUS      CC724960
DEFINITION OGLAB23TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0307C22,
           genomic survey sequence.
ACCESSION  CC724960
VERSION    CC724960.1  GI:32143893
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 867)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
              Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
              Consortium for Maize Genomics
              Unpublished (2002)
              Other GSSs: OGLAB23TH
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TR
              Class: sheared ends.
              Location/Qualifiers
                source
                  1..867
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone="ZMMBma0307C22"
                    /clone_lib="ZM 0.7 1.5_KB"
                    /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                    methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.:      1.6e-114      Length:      867
Score:          1183.00      Matches:    236
Percent Similarity: 90.15%      Conservative: 2
Best Local Similarity: 89.39%      Mismatches: 7
Query Match:    77.78%      Indels:     20
DB:             9      Gaps:         1

US-10-042-894A-8 (1-289) x CC724960 (1-867)

QY    1  MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys  20
      |||||
DB    131  ATGCCCGACCTCCACCCCGGAGACCACCAAGTCCGGCTCACCGGCTCCCGCCAGCAG  190

QY    21  LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg  40
      |||||
DB    191  CTGGGCCCACTCATCGACGACTCTGGCTCTTCTTACAGCCGCTCCAGCGCGGCGACCGT  250
```



```

QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 251 GGGAGCAGCAGGTCGCTCTTAAGAGGCTTCGCCCCACGCCGCCGCTCCGCCGCCG 310
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 311 ATCCGAGACACCTTCTCCCGCGTTCCACGCGCAGCGACTCTCCACCGAGGCGCAG 370
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB 371 CCGCGGAGCGCATCCGACCTCGCTCTGACGACCTCTCCGCGGGTTTGAGGCGCC 430
QY 101 CysValAlaAspTyrLeuGlyAlaLeuThrTrpProProSerProGluProTyr 120
DB 431 TGGCTCGCAGACATCAAGATCGGTGCATCACTG----- 465
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 466 -----ACCACGAGCGTTCTGCTCGGATTCGCGTC 495
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 496 TCCGGCGTCCGAGTCTCGCGCCCGAGGCGCGCTGTGGCGGACGCGCGCGAGGTG 555
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB 556 AAGCTATGAGACATGTCGCGCGTCCCGCGGCTCCGCGGCTACGAGGAGAAAGTGGAGTCT 615
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValle 200
DB 616 CGACGAGGGAGTGAATGCGCGCTCGCGGCGGCTGACGAGGAGAAAGTGGAGTCT 675
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
DB 676 GTCAAGCTGCGCGAGCTCAAGCGTGTTCGAGGGGCGACACTCTGTTCACCTTCTACTC 735
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 736 GCGCTCGATTCTTCGCGCTATGATGCTGTCAGTCGACGAGCGGAGGTGGGGGTGG 795
QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
DB 796 GGTAAACAGTGAAGCTGGTGAATTTGCCCATGTGGCGAGGGGTGATGGGTGATTGACCA 855
QY 260 sAsnPheLeu 263
DB 856 CAACCTTCCTG 865

RESULT 6
CG284572 719 bp DNA linear GSS 25-AUG-2003
LOCUS OGMIF61TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0584K02,
DEFINITION genomic survey sequence.
ACCESSION CG284572
VERSION CG284572.1 GI:34198786
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 719)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Reinick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGMIF61TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

```

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

## FEATURES

## source

```

1..719
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0584K02"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBESK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4,91e-111 Length: 719
Score: 1149.00 Matches: 230
Percent Similarity: 89.92% Conservative: 2
Best Local Similarity: 89.15% Mismatches: 7
Query Match: 75.54% Indels: 20
DB: 9 Gaps: 1

```

US-10-042-894A-8 (1-289) x CG284572 (1-719)

```

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB 3 ATGCCGAGCTCCACCCGCGGAGCACCAAGTCCGCGGTCCCGCGCTCCGCGCAG 62
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 63 CTGGGCCCCACTCATCGACGACTCTGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGT 122
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 123 GGGGAGCAGGAGTCCGCTTCTATGAGGCGTTCCTCCGCCACGCCCGCTCCGCGCCGC 182
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 183 ATCCGAGACACCTTCTTCCCGCGTTCACGCGCACCGGACTCTCTCCACCGAGGCGCAG 242
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB 243 CCGGGGAGCGCGCATCCGACACTCTGCTCTCGACGACTCTCTCGCGGGTTTGAGGCGCC 302
QY 101 CysValAlaAspTyrLeuGlyAlaLeuThrTrpProProSerSerProGluProTyr 120
DB 303 TGGTCCGACAGACATCAAGATCGGTGCATCACTG----- 337
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 338 -----ACCACGAGCGTTCTGCTCGGATTCGCGGTC 367
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 368 TCCGGCGTCCGAGTCTGCGCGCCCGAGGCGCGCTGTGGCGGACGAGCGCGCGAGGTG 427
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB 428 AAGCTATGAGACATTCGCGCGTCCCGCGCTCTCCGCGCTACGCTCATCGCTTGC 487
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValle 200
DB 488 CGACGAGGGAGTGAATGCGCGCTCCGCGCGCGGTGTACGAGGAGAAAGTGGAGTCTT 547
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
DB 548 GTCAAGCTGCGGAGCTCAAGCGGTGTTCGAGGGGCGACACTCTGTTCACCTTCTACTC 607
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 608 GGGCTCGATTCTTCTGGGCTATGATGCTGCTGTCAGTCGACGAGGCGGAGGTGGGGTGG 667

```

```

QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 668 GGTAACAGTGAAGCTGGTGACTTTGCCCATGTGGCGAGGGTGATGGGGTG 719

RESULT 7
CAL34480 801 bp mRNA linear EST 24-SEP-2003
SCJFRT1061H11.g RT1 Saccharum officinarum cDNA clone SCJFRT1061H11
5', mRNA sequence.
CAL34480
CAL34480.1 GI:35021536
EST.
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 801)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 061 row: H column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .801
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFRT1061H11"
/lab_host="DH10B"
/clone_lib="RT1"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:
Pred. No.: 3,24e-108 Length: 801
Score: 1123.00 Matches: 220
Percent Similarity: 89.58% Conservative: 12
Best Local Similarity: 84.94% Mismatches: 24
Query Match: 73.83% Indels: 3
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CAL34480 (1-801)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 31 ATGTCCGACTTCGCGCGGAGACCAAGTCGCGGCACCGGCCTCCGCCAACAG 90

QY 21 LeuGlyProLeuLysAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 91 CTGGGTCCGCTCATCGACGGCTCTGGCTCTTCTACAGCGCTCCAGTCCGAGTCCGGACCGC 150

```

```

QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 151 GGGGAGCAGAGCTCGCTTCTACAGGCGTTCTCCACCACGCGCGCTCCGGCCGC 210

QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 211 ATCCGGGACACCTTCTTCCCGGGTTTCAAGCAGCGACTCTCTCCACACGAGGCGG 270

QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 271 CCAGGGAGCGCATCCGCACCTCTCTCGACGACCTCTCGCGGGGTGGAGGCGCC 330

QY 101 CysValAlaAspIleValIleGlyAlaIleThr-TripProProSerSer-ProGluProT 120
Db 331 TCGTTCGCGCATCAAGATCGCGCCATCAAGTGGCGCGGAGCTCGCGCAGAGCCCT 390

QY 120 YrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgV 140
Db 391 AGTCTCAAGTGCCTGGTCAAGGACCGCGGACACGAGCATTTCTGTCGATTCGCG 450

QY 140 alSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluV 160
Db 451 TCTCCGGGTCCGGTCCGTCGTCGTCGCGAGGCGCGTGTGGCGACTGAGCGCCGAGG 510

QY 160 alLysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValA 180
Db 511 TGAAGGCTTGACACCCGCGCGTCCGCGCGTCTCCGCGCTACGTTTTCATCGTTG 570

QY 180 laAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValL 200
Db 571 CGACGAGGGATGACCTGCGCATCTCGCGCGCGGTGTACGCGGGENAAGGGAGTTT 630

QY 200 euSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrS 220
Db 631 TGTCAAAATTGGCGAGCTTAAGGCGTGTGTAGGAAACAGACTCTGGTCCACCTTTACT 690

QY 220 erAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGly-AspGlyGly 239
Db 691 CGGGGCGGATTTCTTCTGGGCTATGAAGCTAGTGGCTGCAACACCGGAAGGTAAAGGT 750

QY 240 GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256
Db 751 GGGGTGAAGTAAACTGGGGGCCCTTTGCCATGTGGCCCAAGGGAAGG 801

RESULT 8
CAL24555 656 bp mRNA linear EST 25-SEP-2003
SCBGF15081A03.g Saccharum officinarum FL5 Saccharum officinarum
cDNA clone SCBGF15081A03 5', mRNA sequence.
CAL24555
CAL24555.1 GI:35324301
EST.
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 656)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br

```

Plate: 081 row: A column: 03  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1. .656  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCBGL5081A03"  
/lab\_host="DH10B"  
/clone\_lib="Saccharum officinarum FL5"  
/notes="Organ: Developed inflorescence (20cm-long) without rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Developed inflorescence (20cm-long) without rachis]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details source of RNA and library construction can be obtained at http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:  
Pred. No.: 653e-99 Length: 656  
Score: 1034.00 Matches: 196  
Percent Similarity: 97.60% Conservative: 7  
Best Local Similarity: 94.23% Mismatches: 4  
Query Match: 67.98% Indels: 1  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CA245555 (1-656)

Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 33 ATGTCGACCTCCGCCGCGGAGCACCAAGTCCCGGCCACCGCGCTCCGCCAACAA 92  
Qy 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 93 CTGGTCCGCTCATCGACGCTCTGGCTCTTTTACAAAGCGCTCCAGGTCCGGGACCGC 152  
Qy 41 GlyLHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
Db 153 GGGAGCAGCGTCTTCTACAGCGCTTCTCCACCGCGCGCTCCGCCGCGCGC 212  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 213 ATCGGACACCTTCTTCCCGCGGTTTCACGGCACCGGACTCTCTCCACCGCGCGG 272  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 273 CCAGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCGCGGGCTGGAGCGCC 332  
Qy 101 CysValAlaAspIleLysIleGlyValAlaIleThrTrpProSerSerProGluProTyr 120  
Db 333 TGGTCCCGGACATCAAGATCGGCCATCAGTGGCGCGGAGCTCGCGGAGCCCTAC 392  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 393 GTCCCAAGTGCTTCGCCAAGGACCGCGGGACCAACGAGCATTTCTGCTCGGATTCGCG 452  
Qy 141 SerGlyValArgValValGlyProGluValAlaValTrpArgThrGluArgProGluVal 160  
Db 453 TCCGGCTCGGGTCTCGGTCGCGAGCGCGCGCTGTGGCGGACTGAGCGCGCGGAGTG 512  
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180  
Db 513 AAGCCCTGGACACCGCGCGGCTCGCGCGTGTCTCGGCGCTACGTTTCATCGTTGCC 572  
Qy 181 AspGlyMetAspCysAlaLeuAlaAlaValTyrGlyGly-LysGlyGlyValLe 200  
Db 573 GACGAGGGATGGACTCGCACTCGCGCGCGGTGTACGCGCGCAAAAGGGGAGTCTT 632

Qy 200 userGlnLeuArgGluLeuLys 207  
Db 633 GTCACAGCTGCGGAGCTTAAG 654

RESULT 9

CA130685

LOCUS

DEFINITION

5', mRNA sequence.

CA130685.1 GI:35013914

EST.

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

1 (bases 1 to 722)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 004 row: G column: 05

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. .722

/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCCCT1004G05"

/lab\_host="DH10B"

/clone\_lib="RTL"

/notes="Organ: Root tips (0.3cm-long) from adult plants;

Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [Root tips

(0.3cm-long) from adult plants]. cDNA was prepared from

polyA+ mRNA using Superscript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:  
Pred. No.: 1,45e-94 Length: 722  
Score: 993.50 Matches: 202  
Percent Similarity: 92.07% Conservative: 7  
Best Local Similarity: 88.99% Mismatches: 17  
Query Match: 65.32% Indels: 4  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CA130685 (1-722)

Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 31 ATGTCGACCTCGC-CGCGCGGAGCACCAAGTCCCGGCCACCGCGCTCCGCCAACAA 89  
Qy 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 90 CTGGGTCGCTCATCGACGCTCTGGCTCTTCTACAAAGCGCTCCAGGTCCGGGACCGC 149

```
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 150 GGGAGCACGAGCTCGCTTCTACGAGGCGTTCACACCACCGCCGCTCCGCGCCGC 209
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 210 ATCCGGGACACCTTCTTCCCGCGGTTCACGGACGCGACTCTCCACCGAGGGCGG 269
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 270 CCAGGGAGCGGATCCGCAAGTCCGCTCGAGACTCTCCGCGGGCTGGAGGCGCC 329
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 330 TCGGTGCGGACATCAAGATCGGCGCATCAGCTGGCGCGGAGCTCGCGGAGCCCTAC 389
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 390 GTCCGCCAAGTCCGCTCGCAAGACCGGGACACGAGCATTTCTGCTCGGATTCGCGTC 449
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 450 TCCGGGTCCGGTCTGCTGCTCCGAGGCGCGCTGTGGCGACTGAGCGCCCGGAGGTG 509
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla 180
Db 510 AAGGCTCTGACACCGCGCGCTCGCGCGTCTCGCGGCTTACGTTTCATCTCGTTGCC 569
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaValLe 200
Db 570 GACGAGGGATGACTCGGCACTCGCGCGCGGGGAAACCGCGCGGAAAGGGAGTCTT 629
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220
Db 630 GTCACAGTCCCGAGCTTAAGCGTGGTGGTGGAGGA-CAAACTCTGGATCACTT-TACTC 687
Qy 220 rAlaSerIleLeuLeuGly 226
Db 688 GGGGCGGATCTTTTCGGGC 706
```

```
RESULT 10
CG776236 591 bp DNA linear GSS 29-OCT-2003
LOCUS 1123005E08.x1 1123 - RescueMu Grid L Zea mays genomic, genomic
DEFINITION survey sequence.
```

```
ACCESSION CG776236
VERSION 1.0
KEYWORDS GSS.
SOURCE CG776236.1 GI:38034108
ORGANISM Zea mays
```

```
Zealife; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 591)
Walbot, V.
```

```
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
```

```
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1123005 column: 3
Class: transposon-tagged.
Location/Qualifiers
1. .591
```

```
FEATURES
source
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
```

```
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1123 - RescueMu Grid L"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdd.iastate.edu' and follow the links for
'RescueMu.' Grid L was grown in Molokai in 2001. DNA was
extracted from leaf strips, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 1.58e-94 Length: 591
Score: 992.00 Matches: 187
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.47% Mismatches: 0
Query Match: 65.22% Indels: 0
DB: 9 Gaps: 0
```

```
US-10-042-894A-8 (1-289) x CG776236 (1-591)
```

```
Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 28 ATGTCCGAGCTCCACCGCGGAGCACCAGTCCGCGGCACCGCGCTCCGCCAGCAAG 87
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 88 CTGGGCGCGCTCATCGACGGCTCCGCGCTCTTCTTAAAGCCGCTCTCAGCGCGCGCCGT 147
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 148 GGGGAGACGAGGTGCTCTTATGAGGGTTCCTCGGCCACCGCGCTCCGCCCGCC 207
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 208 ATCCGAGACACCTTCTCCCGGTTCCACGCGACGCTCTCCGCCACCGAGCGCGAG 267
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 268 CCGGGGAGCGGCTCTCCTCTCTCGACGACCTCTCTCGGGGGTTTCAGGCGGCC 327
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 328 TCGGTGCGGACATCAAGATCGGCGCATCAGTGGCCACCGAGTTCGCCGAGCGCTAC 387
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 388 ATCGCAAAGTGTCTCGCCAAAGACCGCGGACACGAGCGTTCTGCTCGGATTCGCGCTC 447
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 448 TCGGCGTCCGAGTCTCTCGGCCCGCGCGCTGTGGCGGACCGAGCGCCGAGGGTG 507
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla 180
Db 508 AAGGCCATGGACACCGCGCGGTCCGCGCGTCTCGCGCGCTCTCGCGCGCTACGTGTCCG 567
Qy 181 AspGluGlyMetAspCysAlaLeu 188
Db 568 GACCAGGGGATGGACTGTGCGCTC 591
```

```
RESULT 11
CG202412
LOCUS SCRLFL1009H02.g FL1 Saccharum officinarum cDNA clone SCRLFL1009H02
DEFINITION 5', mRNA sequence.
ACCESSION CA202412
```

```

VERSION          CA202412.1  GI:35237651
KEYWORDS          EST.
SOURCE            Saccharum officinarum
ORGANISM          Saccharum officinarum

REFERENCE
1 (bases 1 to 671)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCBST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 009 row: H column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..671
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCRLFL1009H02"
/lab_host="DH10B"
/clone_lib="FL1"
/note="Organ: Inflorescence at beginning of development
(1cm-long); Vector: pSport1; Site_1: SalI; Site_2: NotI;
An unidirectional cDNA library generated from
[Inflorescence at beginning of development (1cm-long)].
cDNA was prepared from polyA+ mRNA using Superscript
Plasmid System kit (Invitrogen). The double-strand cDNAs
were fractionated in a sepharose CL-2B 40cm-columns and
fragments sizing between 0.8 and 1.5 Kb were
directionally cloned into the vector. Details of each
source of RNA and library construction can be obtained at
http://sucst.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:
Pred. No.: 6,4e-94 Length: 671
Score: 987,00 Matches: 192
Percent Similarity: 91,93% Conservative: 13
Best Local Similarity: 86,10% Mismatches: 17
Query Match: 64,89% Indels: 2
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CA202412 (1-671)

Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Db 2 GGGAGACACAGTCTGCTTCTACGAGCGGTTCACACACACCGCGCGTCCCGCCGC 61
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Db 62 ATCGGGACACCTTCTTCCCGCGGTTCACGGCAGCGACTCTCCACCGAGCGCGG 121
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Db 122 CCAGGGAGCGCATCGACCTCGTCTCGACGACCTCTCGCGGGGCTGGAGGCGCC 181
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Qy 101 CysValAlaAspIleValIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Db 182 TGCCTCGCCGACATCAAGATCGGCGCATCAGTGGCGCGGAGCTCGCGGAGCCCTAC 241
|||||.....:|||||.....:|||||.....:|||||.....:|||||
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
|||||.....:|||||.....:|||||.....:|||||.....:|||||

```

us-10-042-894a-8.p2n.rst

242 GTCCCAAGTGGCTCGCCAAAGGACCGCGGACCCACAGAGATTCTGTCTGGATTCCGCGTC 301

Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTIPArgThrGluArgProGluVal 160

Db 302 TCGGGCTCGGGTCTCGTCCGAGGGCGCGTGTGGCGGACTGAGCGCCCGGAGGTG 361

Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180

Db 362 AAGGCCCTGGACACCGCGCGTCCGCGCGTCTCCGCGCTACGTTTCATCCGTGCC 421

Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200

Db 422 GACGAAGGATGACTGCGCACTCGCGCGCGGTGTACGCGCGGAGGAGGAGTTCG 481

Qy 201 SerGlnLeuArgGluLeuLysAlaTIPheGluGluGlnThrLeuPheHisPheTyrSer 220

Db 482 TCACAGCTCGCGAGCTTAAGGCGTGTTCGAGGAGCAGACTCTGGTCCACTTCTACTCG 541

Qy 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240

Db 542 GCGGTGAATCTTTTGGGCTATGATGCTAGTGCAAGTGACGACGAGCGAGGTGAAGTGG 601

Qy 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp-GlyValIleAspHi 260

Db 602 TTGAAGGTGAA-CTTGTGACTTTTCCCATTTTGTCCAAAGTTGTAGTAGGGGAATGAACCA 660

Qy 260 'eAenPhe 262

Db 661 AATTTTC 667

RESULT 12

AY109355 3374 bp mRNA linear HTC 17-OCT-2002

LOCUS Zea mays CL49\_2 mRNA sequence.

DEFINITION Zea mays CL49\_2 mRNA sequence.

ACCESSION AY109355

VERSION AY109355.1 GI:21213015

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 3374)

Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

Unpublished (2002)

2 (bases 1 to 3374)

Coe,E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

Location/Qualifiers

1..3374

/organism="Zea mays"

/mol\_type="mRNA"

/db\_xref="MaizeDB:632179"

/db\_xref="taxon:4577"

/clone\_lib="Maize Mapping Project/DuPont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize

FEATURES

source

Mapping Project"

ORIGIN

Alignment Scores:  
Pred. No.: 2.63e-92 Length: 3374  
Score: 981.00 Matches: 205  
Percent Similarity: 71.88% Conservative: 2  
Best Local Similarity: 71.18% Mismatches: 62  
Query Match: 64.50% Indels: 20  
DB: 3 Gaps: 1

US-10-042-894A-8 (1-289) x AY109355 (1-3374)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
DB 72 ATGCCCGACTCCACCGCGGAGACCAAGTCGCGGTCACCGCCCTCCGCGACAG 131  
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 132 CTGGGCCCACTCATCGACGACTCTGGCCTCTTCTACAAGCGCTCCAGGCGCGGACCGT 191  
QY 41 GlyLHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 192 GGGGACGACGAGTCCCTTCTATGAGCGGNNNNNNNNNNNNNNNNNNNNNNNN 251  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 252 NTCCGAGACACCTTCTNNNNNGTTCCAGGCGCGGACTCTCTCCACCGAGGCGAG 311  
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
DB 312 CCGGGGAGCGCATCCGCACCTCGCTCTCGACGACTCTCTCGCGGGTTTGAGGCGCC 371  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
DB 372 TGGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 406  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 407 -----ACCACGAGCGTTCGTCTCGATTCGCGTC 436  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
DB 437 TCGCGCGTCCGAGTCTCGCGCCGAGGCGCGGTGNNNNNNNNNNNNNNNNNN 496  
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
DB 497 AAGGCTATGGACATTCTCGCGCGTCCGCGCGTGTCTCCGCGCTACGTGTCTCCGCTTGC 556  
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeu 200  
DB 557 CGACGAGGGATGGACTCGCGCTCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTT 616  
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
DB 617 GTCACAGCTGCGGAGCTCAAGCGGTGTTCGAGGGGCGAGACTNNNNNNNNNNNN 676  
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyAspGlyGlyGly 240  
DB 677 NNN 736  
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260  
DB 737 NNN 796  
QY 260 sAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
DB 797 CAACTTCCTGGGCGGCTCTGCTAGCTGATCAAGTTTGTCTGACATTGTTTCCAGAGAC 856  
QY 280 rProHisThrGlnProLeuGly 287  
DB 857 TCTCTACAGCGACGCTTTGGN 878

RESULT 13

CC724950/c  
LOCUS CC724950 722 bp DNA linear GSS 23-JUN-2003  
DEFINITION OGLAB23TH ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0307C22,  
genomic survey sequence.

ACCESSION CC724950  
VERSION CC724950.1 GI:32143883  
KEYWORDS GSS

SOURCE Zea mays  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 722)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics  
Unpublished (2002)

Other GSSs: OGLAB23TV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208

Email: whitelaw@tigr.org  
Seq primer: IR

Class: sheared ends.

Location/Qualifiers  
1..722

/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="caxon:4577"  
/clone="ZMMBMA0307C22"  
/clone\_lib="ZM 0.7 1.5\_KB"  
/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 1.04e-92 Length: 722  
Score: 976.00 Matches: 197  
Percent Similarity: 88.44% Conservative: 2  
Best Local Similarity: 87.56% Mismatches: 7  
Query Match: 64.17% Indels: 20  
DB: 9 Gaps: 1

US-10-042-894A-8 (1-289) x CC724950 (1-722)

QY 66 PheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluProHis 85  
DB 721 TTCCCCCGGTTCCACGGCACGCGACTCTCTCCACCGAGGCGCAGCCCGGGAGCGCAT 662  
QY 86 ProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIle 105  
DB 661 CGCACCTCTGCTCTCGACGACTCTCTCGGGGGTTGAGGCGCCCTGCGTCGACACATC 602  
QY 106 LysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLysCysLeu 125  
DB 601 AAGATCGGTGCCATCACGTG----- 582  
QY 126 AlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145  
DB 581 -----ACCACGAGCGTTCGTCTCGATTCGCGCTCCGCGGCTCCGAGTC 537  
QY 146 ValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAlaMetAspThr 165  
DB 536 GTCGGCCCCGAGGGGCGCGTGTGGCGACGCGCGCCCGGAGGTGAAGCTATGACATT 477  
QY 166 AlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAlaAspGluGlyMetAs 185  
DB 476 GTCGGCGTCCGCGCGTCTCCGCGCTACGTGTCTATCCGCTTCCGCGAGGGGATGA 417



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Poideae: Triticeae: Triticum.  
1 (bases 1 to 1191)  
Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,  
Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Jarosche,A.,  
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,  
Penniket,C., Roach,J.L. and Sarhan,F.  
Functional Genomics of Abiotic Stress In Wheat and Canola Crops  
Unpublished (2003)  
Contact: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas.ests@usask.ca  
This sequence is the direct result of the Base calling software  
Phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [1,797].  
Plate: L6B002 row: B column: 15.

FEATURES  
source

Location/Qualifiers  
1. .1191  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"  
/notes="Organ: Crown and leaf; Vector: pCMV.SPORTS; Crown  
(50%) and leaf (50%) tissues from wheat cultivar Norstar  
after short exposure times to low temperature in the light  
and in the dark. 12 mRNA populations were combined before  
constructing the library. The first 6 populations: After 7  
days of growth at 20°Cs from wheat cultivar Norstar after  
short exposure times to low temperature in the light and  
in the dark. 12 mRNA populations were combined before  
constructing the library. The first 6 populations: After 7  
days of growth at 20, wheat plants were transferred to 4C  
in the light. 1cm crown sections and green leaf tissue were  
separately harvested after 1, 3, and 6 hours of low  
temperature exposure. The last 6 populations: After 7 days  
of growth at 20C, wheat plants were transferred to 4C in  
the dark. 1cm crown sections and green leaf tissue were  
separately harvested after 1, 3, and 6 hours of low  
temperature exposure. First strand synthesis in this  
library was done in the presence of methylated dCTP  
thereby protecting from internal cleavage with NotI. In  
addition, this library used a primer for second strand  
synthesis that annealed to an artificial sequence (RNA  
oligo) added before first strand synthesis. Therefore when  
sequences from EST generated from this library will be  
masked for vector and adaptor sequences, an additional  
masking step will have to be included to mask this RNA  
oligo that is common to all clones (sequence  
CGACTGGACACGAGGACACTGACATGGACTGAGGAGTAGAAA)."

## ORIGIN

Alignment Scores:  
Pred. No.: 4,11e-87 Length: 1191  
Score: 926.00 Matches: 183  
Percent Similarity: 73.63% Conservative: 18  
Best Local Similarity: 67.03% Mismatches: 68  
Query Match: 60.88% Indels: 4  
DB: 7 Gaps: 2

US-10-042-894A-8 (1-289) x CK211413 (1-1191)

Qy 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIysLeuGlyProLeuIleAsp 26  
Db 1064 CCGGAGACACGAGTCCCGGCCACCGCGCGCCAGAAAGATTGGCCCTTTGTGAC 1005	
Qy 27 GlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArgGlyHisGluValAla 46	

Db 1004 AGCCCGGCCCTTTTACAAAGCCGTTGAAGGCCCTTAGCCCGCGGACAGAGGTTCCCC 945  
Qy 47 PheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAspThrPhePhe 66  
Db 944 TTTAACACGGCCCTTTTCGCCCCACCCGAAAGTTCCGCCCGCATCCGGACACCTTCTTC 885	
Qy 67 ProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisPro 86	
Db 884 CGGCGCTGACACGGCAACGCCATTTCACCAGCGGCTCTTCCCGCGGAGTCCACACCC 825	
Qy 87 HisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIleIys 106	
Db 824 CACCTCATCTGGACGACCTCTCTCAAGGGCTTCGCCGCGCCCTCCGTCAACCGACATCAAG 765	
Qy 107 IleGlyAlaIleThrTrpProSerSerProGluProTyrIleAlaIysCysLeuAla 126	
Db 764 ATCGCGCGCTGCACGTGGCCCGCGCGCGAGCCCTAGCTCACCAAGTGCCTCGCC 705	
Qy 127 LysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValVal 146	
Db 704 AAGGACCGCGGCTCCACGAGGTCTCTCGGCTTCCGGGTCTCGGGGTCTATGTATCC 645	
Qy 147 GlyProGluGlyAlaValTrpArgThrGluArgProGluValIysAlaMetAspThrAla 166	
Db 644 GACGCCAGCGCGCGCTCTGGCGCCGACAGGTCCGAGTCAAGGGGACGACATCCCC 585	
Qy 167 GlyValArgValLeuArgArgTyrValSerSerValAlaAspGlu-----GlyMet 184	
Db 584 GCGGTGCGCGCATCTCCGCCGTACGTGTCTCGTCCGCGCGGACGCGGGGCGGAG 525	
Qy 185 AspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArg 204	
Db 524 GACTCGCGCTCGCGCGCGCGCTGTACGGGGCGAGGGGGGTCTGGCGGAGTGGCC 465	
Qy 205 GluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeu 224	
Db 464 GAGCTCAAGGGGTGTTCGAGGTGCAGACGCTGTTCCACTTCTACTCGCGCTCGGTGCTG 405	
Qy 225 LeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValIys 244	
Db 404 CTGAGCTACGACGCCCAACGCGGTGACCGCG-----CCGCGCGCGCGCGCGAGGTGAAG 351	
Qy 245 LeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGly 264	
Db 350 CTGGTGGACTTCGCCCATGTGTGGAGAGCGAGGGGTGTATCATCACCACTTCTTGGGC 291	
Qy 265 GlyLeuCysSerLeuIleLysPheValSerAspIleVal 277	
Db 290 GGCCTCTGCTCCTCATCAAGTTTCATAGACGACATTGTC 252	

Search completed: June 17, 2005, 06:34:53  
Job time : 3703 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 05:35:00 ; Search time 2130 Seconds  
(without alignments)  
842.244 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111952\_8386/app\_query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOPCLI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=ptp -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10042894 @cgn 1 1 480 @runat\_15062005\_111952\_8386  
-NCFU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1521	100.0	1344	14	US-10-042-894A-7	Sequence 7, Appl
2	1485	97.6	923	14	US-10-042-894A-5	Sequence 5, Appl
3	1484	97.6	1426	20	US-10-425-115-74438	Sequence 74438, A
4	1435	94.3	923	14	US-10-042-894A-3	Sequence 3, Appl
5	1406.5	92.5	1169	14	US-10-042-894A-1	Sequence 1, Appl
6	1307	85.9	3416	14	US-10-042-894A-20	Sequence 20, Appl
7	1233	81.1	899	14	US-10-042-894A-15	Sequence 15, Appl
8	1148.5	75.5	1108	19	US-10-437-963-4650	Sequence 4650, Ap
9	847.5	55.7	696	19	US-10-437-963-4649	Sequence 4649, Ap
10	808.5	53.2	643	14	US-10-042-894A-17	Sequence 17, Appl
11	742.5	48.8	1020	14	US-10-042-894A-13	Sequence 13, Appl
12	722.5	47.5	1195	14	US-10-042-894A-11	Sequence 11, Appl
13	644.5	42.4	1105	14	US-10-042-894A-9	Sequence 9, Appl
14	644.5	42.4	1345	18	US-10-424-599-106467	Sequence 106467,
15	609	40.0	776	20	US-10-425-115-74436	Sequence 74436, A
16	464	30.5	519	14	US-10-042-894A-18	Sequence 18, Appl
17	329	21.6	353	14	US-10-042-894A-19	Sequence 19, Appl
18	306.5	20.2	484	10	US-09-770-961-829	Sequence 829, App
19	273.5	18.0	464	9	US-09-770-444-340	Sequence 340, App
20	253	16.6	295	20	US-10-425-115-10743	Sequence 10743, A
21	201.5	13.2	2212	20	US-10-478-146-20	Sequence 20, Appl
22	193	12.7	1341	18	US-10-641-643-809	Sequence 809, App
23	192	12.6	2890	18	US-10-415-011-37	Sequence 37, Appl
24	191	12.6	2608	17	US-10-094-749-1021	Sequence 1021, Ap
25	191	12.6	2608	21	US-10-887-553A-1159	Sequence 1159, Ap
26	189	12.4	1737	9	US-09-731-872-69	Sequence 69, Appl
27	189	12.4	1737	10	US-09-876-997-69	Sequence 69, Appl
28	189	12.4	1737	21	US-10-643-836-69	Sequence 69, Appl
29	189	12.4	1737	21	US-10-956-157-2095	Sequence 2095, Ap
30	189	12.4	1748	15	US-10-037-270-712	Sequence 712, App
31	189	12.4	1748	17	US-10-117-722-712	Sequence 712, App
32	189	12.4	1757	9	US-09-731-872-76	Sequence 76, Appl
33	189	12.4	1757	10	US-09-876-997-76	Sequence 76, Appl
34	189	12.4	1757	21	US-10-643-836-76	Sequence 76, Appl
35	189	12.4	1791	14	US-10-198-846-10945	Sequence 10945, A
36	189	12.4	1990	9	US-09-731-872-174	Sequence 174, App
37	189	12.4	1990	20	US-09-876-997-174	Sequence 174, App
38	189	12.4	1990	21	US-10-643-836-174	Sequence 174, App
39	182	12.0	1863	17	US-10-120-988-103	Sequence 103, App
40	178.5	11.7	1893	17	US-10-120-988-102	Sequence 102, App
41	168	11.0	4461	19	US-10-618-941-57	Sequence 57, Appl
42	152.5	10.0	546	19	US-10-437-963-5687	Sequence 5687, Ap
43	145.5	9.6	1782	9	US-09-954-456-87	Sequence 87, Appl
44	145.5	9.6	1782	9	US-09-954-456-722	Sequence 722, App
45	145.5	9.6	1782	14	US-10-175-523-196	Sequence 196, App

ALIGNMENTS

RESULT 1

US-10-042-894A-7  
; Sequence 7, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shih, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465

```
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(921)
US-10-042-894A-7

Alignment Scores:
Pred. No.: 4,71e-172 Length: 1344
Score: 1521.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-7 (1-1344)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
DB 52 ATGTCGACCTCCACCCTCCGCGGAGCACCAGTCGCGGCCACCGCGCTCCGCGCAGCAAG 111
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 112 CTGGGCGCGCTCATCGACGCGCTCCGCGCTCTTCTACAAGCGCTCCAGGCGCGCGACCGT 171
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 172 GGGGACACGAGGTCCCTCTATGAGGCTTCTCCGCCACCGCCGCTCCGCGCGCC 231
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 232 ATCCGAGACACCTTCTCCCGCGTTCCAGGCGCGACTCTCTCCACCGCGCGCAG 291
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
DB 292 CCGCGGAGCGCATCTCCACCTCGCTCCGACGACTCTCTCGCGGGTTTCAGGCGCC 351
QY 101 CysValAlaAspIleValIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 352 TGGTTCGACAGACATCAAGATCGCGCCCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 411
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 412 ATGCCCAAGTGCCTGCCAAGGACCGCGGACCACGAGCGTTCTGCTCGGATTCGCGCTC 471
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
DB 472 TCGGCGCTCCGAGTCTCGCGCCCGAGGCGCGCTGTGGCGGACGAGCGCGCGAGGTG 531
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerValAla 180
DB 532 AAGGCAATGACACCCCGCGCGTCCCGCGCTCTCCGCGCTACGTGTATCCGTTCGC 591
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200
DB 592 GACGAGGGATGAGTGTGCGCTCGCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTG 651
QY 201 SerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSer 220
DB 652 TCACAGCTGCGGAGCTCAAGGCGTGGTTCGAGGAGCAGACTCTGTTCACATTTCTACTCG 711
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 712 GCGTCATTTCTTGGGCTATGATGCTGTGCGAGTGCAGACGCGCGAGATGGGGTGG 771
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 772 GTGACGGTGAAGCTGTGGACTTTGCCCATGTGGCCGAGGGTGTGTTGATTTGACCAC 831
```

```
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
DB 832 AACTTCTCTGGCGCGCTCTGCTCGTGATCAAGTTTCTTTCGATTGTTCCGGAGACT 891
QY 281 ProHisThrGlnProLeuGlyProSer 289
DB 892 CCTCATACGACGCTTTGGTCTTCT 918

RESULT 2
US-10-042-894A-5
; Sequence 5, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(922)
US-10-042-894A-5

Alignment Scores:
Pred. No.: 6e-168 Length: 923
Score: 1485.00 Matches: 283
Percent Similarity: 98.27% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 5
Query Match: 97.63% Indels: 0
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-5 (1-923)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
DB 53 ATGCCCGACCTCCACCGCGCGGAGCACCAGTCGCGGTCACCGCGCTCCGCGCAGCAAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CCGGCGCGCTCATCGACGCGTCCGCGCTCTTCTACAAGCGCTCCAGGCGCGCGACCGT 172
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 173 GGGGAGCAGAGTTCGCTTCTATGAGCGTTCCTCCGCCACCGCGCTCCGCGCGCC 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTTCCCGCGTTCACGCGCGCGACTCTCTCCACCGAGGCGCAG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGGCGGAGCGCATCCGACCTCGCTCTCGACGACCTCTCTCCGCGGATTTGAGGCGCC 352
QY 101 CysValAlaAspIleValIleGlyAlaIleThrTyrProProSerSerProGluProTyr 120
DB 353 TCGCTCGGAGACATCAAGATCGCGCCATCAGTGGCCACCGAGTTCGCGGAGGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
```

```
Db 413 ATGCCCAAGTGCCTCGCCATGACCGCGGACACGAGCGTTCTGCTCGGATTCGCGTC 472
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrrArgThrGluArgProGluVal 160
Db 473 TCCGGGTCCGAGTCGTGCGCCCGAGGGCGCGTGTGGCGGACGGAGCGCCCGGAGGTG 532
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
Db 533 AAGGCCATGGACACCGCGCGGTCCGCGGTGCTCGGGGCTACGTGTCTATCGTTGCC 592
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyLysGlyValLeu 200
Db 593 GACGAGGGATGAGTCTGCGCTCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTG 652
Qy 201 SerGlnLeuArgGlnLeuLysAlaTrrPheGluGlnThrLeuPheHisPheTyrSer 220
Db 653 TCACAGCTGCGGAGCTCAAGGCGGTGTTCGAGGAGCAGACTCTGTTCCTACTCG 712
Qy 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db 713 GCGTCGATTCCTTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
Qy 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
Db 773 GTACCGGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGGTGATGGGGTGATTGACCAC 832
Qy 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
Db 833 AACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACATTGTTCCAGAGACT 892
Qy 281 ProHisThrGlnProLeuGlyProSer 289
Db 893 CCTCAGACGAGCCTTTGGGTCTCTCT 919

RESULT 3
US-10-425-115-74438
; Sequence 74438, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 74438
; LENGTH: 1426
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167886C.1
US-10-425-115-74438

Alignment Scores:
Pred. No.: 1 39e-167 Length: 1426
Score: 1484.00 Matches: 286
Percent Similarity: 98.96% Conservative: 0
Best Local Similarity: 98.96% Mismatches: 3
Query Match: 97.57% Indels: 1
DB: 20 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-425-115-74438 (1-1426)
Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 120 ATGTCCGACCTCCACCGCGCGGAGCACCAAGTCGCCCGCCACCGCGCTCCGCGAGCAAG 179
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
```

```
Db 180 CTGGGCGCGCTCATCGACGCGCTCCGCGCTCTTCTACAAGCGCTCCAGCGCGCGGACCGT 239
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 240 GGGGAGCAGAGGTCCGCTTCTATGAGGCGTTCCTCGCCACCGCCGCTCCCGGCGCGC 299
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 300 ATCCGAGACACCTTCTTCCCGGTTCCACGCGCAGCGACTCTCTCCACCGAGGCGCAG 359
Qy 81 ProGlyGlnProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db 360 CCGGGGAGCGCATCTCACCTCGCTCGACGACTCTCTCGCGGGGTTTCAGGCGGCC 419
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrrProProSerSerProGluProTyr 120
Db 420 TGGCTCGCAGACATCAAGATCGCGCCATCACTGTGGCCACCGAGTTTCGCGGAGCCCTAC 479
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 480 ATCCCAAGTGCCTCGCCCAAGGACCGCGGACCACGAGCGTTCGTCTCGGATTCGCGTC 539
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrrArgThrGluArgProGluVal 160
Db 540 TCCGGGTCCGAGTCTCGCGCCCGAGGGCGCGCTGTGGCGGACGGAGCGCCGAGGTTG 599
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSerValAla 180
Db 600 AAGCCATCGACACCGCGCGGTCCCGCGGTCTCCGCGCTACGTGTCTCATTCGTTGCC 659
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyLysGlyGlyValLeu 200
Db 660 GACGAGGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTTG 719
Qy 201 SerGlnLeuArgGlnLeuLysAlaTrrPheGluGlnThrLeuPheHisPheTyrSer 220
Db 720 TCACAGCTGCGGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTGTCCACTTCTACTCG 779
Qy 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db 780 GCGTCGATTCCTTGGGCTATGATGCTGTTCAGTCCA-CCGCCGAGATGGGGTGGG 838
Qy 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
Db 839 GTGACCGGTGAAGCTGTGGACTTTGCCCATGTGCGCCGAGGGTGTATGGGTGATTGACCAC 898
Qy 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
Db 899 AACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACATTGTTCCGAGACT 958
Qy 281 ProHisThrGlnProLeuGlyProSer 289
Db 959 CCTCATACGACGCTTTGGGTCTCTCT 985

RESULT 4
US-10-042-894A-3
; Sequence 3, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(736)
US-10-042-894A-3

Alignment Scores:
Pred. No.: 5,97e-162 Length: 923
Score: 1435.00 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 94.35% Indels: 1
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-3 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys 20
Db 53 ATCCCGAGCTTCACCGCGGAGCACCAGTCGCGCGTACCGCGCTCCCGCAGCAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 113 CCGGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGCGCGGACCGT 172
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 173 GGGGAGCAGAGTGCCTTTCTATAGGGGTTCCTCCGCCACGCGCGCTCCCGCGCCG 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 233 ATCCGAGACACCTTCTCCCGCGTTCCAGGACCGGACTCTCCCAACCGAGGCGCAG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 293 CCGGGGAGCGCATCCGACCTCTCTCTCGAGACCTCTCCGGGATTTGAGGGCGCC 352
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 353 TGGCTCGCAGACATCAAGATCGCGCGCATCATCGTGCCACCGAGTTTCGCGGAGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 413 ATGCCAAGTGCTTCGCCATGACCGCGGAGCACAGCGTTCGTCTCGATTCGCGGTC 472
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 473 TCCGGGTCGAGTCTGCTCCCGAGGGGCGCGTGTGGCGAGCGCGCGGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAla 180
Db 533 AAGGCTATGGACACCGCTCGCGCGTCCCGCGTGTCTCCGCGTACGTGTATCCGCTTC 592
QY 180 aAspGluGlyMetAspCysAlaLeuAlaIleAlaValTyrGlyLysGlyGlyValle 200
Db 593 CGACGAGGGGAGTGAAGTCTCGCGCGTCCCGCGCGGTGTACGAGGAAAGGTGGAGTCTT 652
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 653 GTCACTGCTCGCGAGCTCAAGGCGTGTTCGAGGAGCAGCCTCTGTCTCACTTCTACTC 712
QY 220 rIlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGly 240
Db 713 GCGCTCGATTCTTCGGGTATGATCTCTGTCAGTCGACGAGCGAGGCGGAGGTGGGGTGG 772
QY 240 vValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 773 GGTAAACAGTGAAGCTGGTGACTTTCCCATATGGCCGAGGGGTGATGGGGTGATGACCA 832
QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
```

```
Db 833 CAACCTTCCTGGCGGGCTCTGCTAGCTGATCAAGTTGTTTTCTGACATGTTTCCAGAGAC 892
QY 280 rProHisThrGlnProLeuGlyProSer 289
Db 893 TCCTCAGACGCGACCTTTGGGTCTTCT 920

RESULT 5
US-10-042-894A-1
; Sequence 1, Application US/10042894A
; Publication NO. US2003009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US2003009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(806)
US-10-042-894A-1

Alignment Scores:
Pred. No.: 2,13e-158 Length: 1169
Score: 1406.50 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 92.47% Indels: 2
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-1 (1-1169)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys 20
Db 84 ATGCCCGACCTTCACCGCGGAGCACCAGTCGCGGTTCACCGCGCTCCCGCAGCAG 143
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 144 CTGGGCGCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGCGCGGACCGT 203
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 204 GGGGAGCAGAGTGCCTTCTATGAGCGTTCCTCCGCCACGCGCGCTCCCGCGCCG 263
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 264 ATCCGAGACACCTTCTCCCGCGGTTCACGCGACGCGACTCTCCCGCAGGCGCAG 323
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 324 CCGGGGAGCGCATCCGACCTCTGCTCTCGACGACTCTCTCGCGGGGTTTGAGGGGCC 383
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 384 TCGCTCGCAGACATCAAGATCGGCGCATCATCGTGCCACCGAGTTCGCGGAGCCCTAC 443
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 444 ATCCCAAGTACCTCGCAAGGACCGCGGAGCCAGAGCGTTCCTCGGATTCGCGGTC 503
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
```

Db 504 TTGC--GTCCAGATCGTCGCGCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGAGGTG 561  
Qy 161 LyAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
Db 562 AAGGCTATGGACACCGTCGCGCGTCCGCGCGTGTCCGCGGCTACGTTGTCATCCGCTGC 621  
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyValLeu 200  
Db 622 CGACGAGGGATGAGACTGCGCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGAGTCTT 681  
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220  
Db 682 GTACAGCTCGCGGAGCTCAAGCATGTTGGAGGACGACACTGTTCCACTTCTACTC 741  
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
Db 742 GCGCTCGATTCTTCTGGGCTATGATGCTGTCAGTCGCAGCGAGCGAGGTGGGGTGG 801  
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyValIleAspHi 260  
Db 802 GGTAAACAGTAGAGCTGTGAGCTTTGCCCCATGTGCGCGGAGGTGATGGGTGATTGACCA 861  
Qy 260 sAsnPheLeuGlyGlyCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
Db 862 CNACTTCTCGCGGAGCTCTGTAGCTGATCAAGTTCGTTCTGACATTGTTCCAGAGAC 921  
Qy 280 rProHisThrGlnProLeuGlyProSer 289  
Db 922 TCCTTAGACGAGCCTTTGGGTCTCTTCT 949

RESULT 6

US-10-042-894A-20  
; Sequence 20, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafaleki, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; FILE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 3416  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (72)...(407)  
US-10-042-894A-20

Alignment Scores:  
Pred. No.: 7,43e-146 Length: 3416  
Score: 1307.00 Matches: 260  
Percent Similarity: 90.34% Conservative: 2  
Best Local Similarity: 89.66% Mismatches: 9  
Query Match: 85.93% Indels: 20  
DB: 14 Gaps: 1

US-10-042-894A-8 (1-289) x US-10-042-894A-20 (1-3416)

Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 72 ATGCCCCGACCTCCACCCCGCGGACCAAGTCCCGGTCTACCCGCGCTCCGCGCAGCAAG 131

Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 132 CTGGGCCCCACTCATCGACGACTCTGGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGT 191  
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
Db 192 GGGGAGCAGAGGTCCGCTTCTATGAGGGTTCCTCCGCCACCGCCGCTCCGCGGCGCC 251  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 252 ATCCGAGACACCTTCTTCCCCCGTTCCAGCGACCGGACTCTCTCCACCGAGGCGCAG 311  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
Db 312 CCCGGGAGCGCATCCGACACTCTCTCGACGACTCTCTCGCGGGGTTTGAGGCGCCC 371  
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 372 TGGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 406  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 407 -----ACCACGAGCGTTCTGCTCGGATTCGCGGTC 436  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 437 TCCGGCGTCCGAGTCTGCGCCCGAGGCGCGTGTGGCGGACGAGCGCCGAGGTG 496  
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
Db 497 AAGGCTATGACATTGTCGCGCGTCCGCGCGTCTCCGCGCTACGTTGTCATCGCTTCG 556  
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200  
Db 557 CGACGAGGGGATGAGCTGCGCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTT 616  
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220  
Db 617 GTACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGGGCGACACTCTGTTCCACTTCTACTC 676  
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyG 240  
Db 677 GCGCTCGATTCTTCTGGGCTATGATGCTGTCAGTCGCACGAGCGGAGGTGGGGGTGG 736  
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
Db 737 GGTAAACAGTAGAGCTGTGAGCTTTGCCCATGTGCGCGGAGGTGATGGGTGATTGACCA 796  
Qy 260 sAsnPheLeuGlyGlyCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
Db 797 CNACTTCTGGGCGGCTCTGTAGCTGATCAAGTTGTTTCTGACATTGTTCCAGAGAC 856  
Qy 280 rProHisThrGlnProLeuGlyProSer 289  
Db 857 TCCTCAGACGAGCCTTTGGGTCTCTTCT 884

RESULT 7

US-10-042-894A-15  
; Sequence 15, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafaleki, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; FILE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12

```
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 899
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(424)
US-10-042-894A-15

Alignment Scores:
Pred. No.: 9.8e-138 Length: 899
Score: 1233.00 Matches: 248
Percent Similarity: 88.42% Conservatives: 4
Best Local Similarity: 87.02% Mismatches: 14
Query Match: 81.07% Indels: 20
DB: 14 Gaps: 1

US-10-042-894A-8 (1-289) x US-10-042-894A-15 (1-899)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB 89 ATGCCGACCTCCACCGCGGAGCACCAGTCCCGGTCCCGCGCTCCCGCAGCAAG 148
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 149 CTGGGCCCCATCATCAGCGCTCTGCGCTCTCTACAGCGCTCCAGCGCGCGGCGCT 208
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 209 GGGGAGCAGCGGTCTCTATGAGCGTCTCCGCCACCGCGCGCTCCCGCGCGC 268
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 269 ATCCGAGACACCTTCTCCCGCGTTCACGCGCAGCTCTCCGCCACCGAGCGCGCAG 328
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 329 CCCGGGAGCGGATCCGATCTCTGCTCTGCGACCTCTCCGGGGTTTGGGGCGCC 388
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB 389 TCGCTCGCAGACATCAAGATCGGTGCTCATCGTGACCATG-AGCGAT----- 435
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 436 -----CTGCTCGGATTCCACGTC 453
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 454 TCCGGGTCCGAGTCTGCGGCCCGAGGGCGCGTGTGGCGAGCGAGCGCGCTTAGGGT 513
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB 514 AAGGCTATGGACATTGTCGGGTCCCGCGGTCTCCGGGTGTCATGTCATCCGCTTGC 573
QY 180 aAspGluGlyMetAspCysAlaLeuAlaValValTyrGlyGlyGlyGlyVal 200
DB 574 CGCGAGGGGATGAGTATGCGCGCTCCGGCGCGGTGTACGGAGGAAAAGTGGAGTCTT 633
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
DB 634 GTCACAGCTGCGGAGCTCAGGCGTGTTCAGGGGCGAGTCTCTTCCATCTTACTTC 693
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyGlyGlyGly 240
DB 694 GCGTTCGATTCTCTGGGTATGATCTCTGAGTCCGAGCAGCGAGGTTGGGGTGG 753
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
DB 754 GGTAAAGTGAAGTGGTGGACCTTGTCCCATGTGGCCGAGGGTGTATGGGGTGATTGACCA 813
```

```
QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
DB 814 CAACCTTCCTGGCGGGCTCTGCTAGCTGATCAAGTTTGTCTGACATTTGTTCCAGAGAC 873
QY 280 rProHisThrGln 284
DB 874 TCCTTAGACGCGAG 886

RESULT 8
US-10-437-963-4650
; Sequence 4650, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 4650
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11514C.1
US-10-437-963-4650

Alignment Scores:
Pred. No.: 1.76e-127 Length: 1108
Score: 1148.50 Matches: 218
Percent Similarity: 84.45% Conservatives: 21
Best Local Similarity: 77.03% Mismatches: 39
Query Match: 75.51% Indels: 5
DB: 19 Gaps: 3

US-10-042-894A-8 (1-289) x US-10-437-963-4650 (1-1108)

QY 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLysLeu 21
DB 7 TCCGACCTCGCGCGCGCGAGCACCAGGTGGCGGGGCACCGCGCTCCCGCCGACAGCTG 66
QY 22 GlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGly 41
DB 67 GCGCGCTCTCTCGAGCGGCGGGGTCTTCTACAAGCCCTCCAGCGCGGGAGCGCGG 126
QY 42 GluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIle 61
DB 127 GAGCAGAGGCGCGCTTCTACCGCGGTTCACCGCGCACCGCGCGTCCCGCCCGGGTTC 186
QY 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81
DB 187 CGGGCGCGCTTCTTCCCGCTTCCACGCGCACCGCGCTCTCCCGCGCGCCAGCGCC 246
QY 82 ---GlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
DB 247 GCGCGCGCGCTTACCGGACATCTCTCGACGACCTCTCCCGCGGCTCTCCCGTCC 306
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB 307 TCGTCCGCGAGCTCAAGATCGCGCTTCACGTGCGCGCGCGATCCCGGAGCCCTAC 366
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 367 GTCGCAAGTGTCTCGCAAGGACCGAGACACACGAGCGCTCTCTCGGCTTCCGCGTC 426
```

```
Qy 141 SerGlyValArgValValGlyProGlu---GlyAlaValTirArgThrGluArgProGlu 159
Db 427 TCCGGCGTCCGGGTGGTTCGATGCCCGGGGGCGCGCGTGTGGCGCCGGACCGGTCCGAG 486
Qy 160 ValLysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerVal 179
Db 487 CTGAAGGGGATCGACGCCCGCGGGTCCGCGCGGTGTCCGCGGCTACGTGTCCACGGGC 546
Qy 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyVal 199
Db 547 GCGCGGACCGGCTCGAGCTCGCGCTCGCGCGCGGTGTACGAGGGGCGGGCGGTC 606
Qy 200 LeuSerGlnLeuArgGluLeuLysAlaTirPheGluGlnThrLeuPheHisPheTyr 219
Db 607 CTGGCTCAGTTCGGGAGCTCAAGGCTGGTTCGAGGAGGCAACCTGTACCTTCTAC 666
Qy 220 SerAlaSerIleLeuLeuGlyTyrAsp-----AlaAlaValAlaAlaGlyGly 236
Db 667 TCGCGGTTCGATTTCTGTCGGCTACGACGCCAATCGCGCGCGCGGTGTCTCCCGAGGT 726
Qy 237 AspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyGly 256
Db 727 GGAAGCGCGCGGTGAAGGTGAAGCTGGTGGATTCGCGCATGTTCGACATGGGGCGGG 786
Qy 257 ValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276
Db 787 GTGATTGACCAACACTTCTTGGCGGGGTCTGCTCGCTCATCAAGTTTCATCGCGGACATT 846
Qy 277 ValProGlu 279
Db 847 GTCGCGGAG 855
```

## RESULT 9

```
US-10-437-963-4649
; Sequence 4649, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 4649
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11513C.1
US-10-437-963-4649
```

## Alignment Scores:

Pred. No.:	1,22e-91	Length:	696
Score:	847.50	Matches:	173
Percent Similarity:	66.31%	Conservative:	14
Best Local Similarity:	61.35%	Mismatches:	28
Query Match:	55.72%	Indels:	67
DB:	19	Gaps:	3

US-10-042-894A-8 (1-289) x US-10-437-963-4649 (1-696)

```
Qy 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLeu 21
Db 7 TCCGACCTCGCCCGCGGACCAAGGTGGCGGGGACCGCGGTCCGCGGACCAAGCTG 66
```

```
Qy 22 GlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGly 41
Db 67 GGGCCGCTCGTCGACGAGGAGGGCTCTTCTACAGCCCTCCAGGCGCGGAGCGGG 126
Qy 42 GluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIle 61
Db 127 GAGCAGGAGCGCGCTTTTACGCGCGGTTCACCGCGCACCGCGCGTCCCGCCCGG 183
Qy 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81
Db 183 ----- 183
Qy 82 GlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaProCys 101
Db 183 ----- 183
Qy 102 ValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIle 121
Db 183 ----- 183
Qy 122 AlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSer 141
Db 184 -----TGCTTCGCCAAGGAGCCCGAGACCCAGCGCGCTCTCGGCTTCGCGTCTCC 237
Qy 142 GlyValArgValValGlyProGlu---GlyAlaValTirPargThrGluArgProGluVal 160
Db 238 GGGTTCGGGTGGTTCGATGCCCGGGGCGCGCGTGTGGCGCCGAGCCGTCGGAGCTG 297
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
Db 298 AAGGATCGACGCGCGCGGTCCGCGCGCTCTCGCGCTACGTGTCCACGGCGGC 357
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200
Db 358 GCGACGCGCTGACTGCGCGCTCGCGCGCGGTGTACGAGGGGAGGGCGCGCTCTG 417
Qy 201 SerGlnLeuArgGluLeuLysAlaTirPheGluGlnThrLeuPheHisPheTyrSer 220
Db 418 GCTCAGCTCGGGAGCTCAAGGCGTGTTCGAGGAGCAACCTGTACCACTTCTACTCG 477
Qy 221 AlaSerIleLeuLeuGlyTyrAsp-----AlaAlaValAlaAlaGlyGlyAsp 237
Db 478 GCGTCGATTCTGTTCGGCTACGACGCCAATGCGCGCGCGCGCTGTCTCCCGAGGTGA 537
Qy 238 GlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 538 AGCGCGGTGTAAAGGTGAAGCTGGTGGACTTCGCGCATGTCCAGATGGGAGCGGGTG 597
Qy 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 598 ATTGACCACAACTTCTTGGCGGGCTCTGCTCGCTCATCAAGTTTCATCGCGGACATTGTC 657
Qy 278 ProGlu 279
Db 658 GCGGAG 663
RESULT 10
US-10-042-894A-17
; Sequence 17, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
```

```
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(643)
; OTHER INFORMATION: n = A, T, C or G
US-10-042-894A-17
```

```
Alignment Scores:
Pred. No.: 5,26e-87 Length: 643
Score: 806.50 Matches: 162
Percent Similarity: 93.14% Conservative: 1
Best Local Similarity: 92.57% Mismatches: 10
Query Match: 53.16% Indels: 3
DB: 14 Gaps: 0
```

US-10-042-894A-8 (1-289) x US-10-042-894A-17 (1-643)

```
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
|||
Db 117 ATCCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTCACCGCGCTCCGCGCAAG 176
|||
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
|||
Db 177 CTGGGCGCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGGCTCCAGCGCGCGACCGT 236
|||
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
|||
Db 237 GGGGACACGAGGTGCGCTCTTATGAGGGTCTCCGCCACCGCGCTCCGCGCGCG 296
|||
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
|||
Db 297 ATCCGAGACACCTTCTTCCCGCGTTCCACGGCAGCGACTCTCCGCCACCGAGGGCGAG 356
|||
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
|||
Db 357 CCGGGGAGCGGATCCGACCTCGCTTCGACGACCTCTCCGCGGGTTGAGGGCGCC 416
|||
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyr 120
|||
Db 417 TGGGTGCGACACATCAAGATCGCGGCATCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 476
|||
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
|||
Db 477 ATCGNCAAGTACCTNGCCAAAGGACCGCGGACCCACGAGCGTTCTGCTCGGATTCGCGTC 536
|||
QY 141 SerGlyValArgValGlyProGluGlyAlaValThrArgThrGluArg-ProGluVa 160
|||
Db 537 TTGC--GTCCGAGTCGTCCGCGCGGCGCGCTGTGGCGGACGAGCGCGCGGGT 594
|||
QY 160 lLysAlaMetAspThr-AlaGlyValArgValLeuArg 173
|||
Db 595 GAANGCTATGGACCCCGTCGGNGNCCGCGGTGTTTCGG 635
|||
```

## RESULT 11

```
US-10-042-894A-13
; Sequence 13, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
```

```
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Parthenium argentatum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(908)
US-10-042-894A-13
```

```
Alignment Scores:
Pred. No.: 7,93e-79 Length: 1020
Score: 742.50 Matches: 149
Percent Similarity: 67.83% Conservative: 45
Best Local Similarity: 52.10% Mismatches: 79
Query Match: 48.82% Indels: 13
DB: 14 Gaps: 5
```

US-10-042-894A-8 (1-289) x US-10-042-894A-13 (1-1020)

```
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
|||
Db 24 CTCAAGGCCCCAGATCATCAGGTTCGTGACATGAAGCTGGGCTCGGGAAGCTTGGCCCA 83
|||
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
|||
Db 84 CTCATTGATGATTCAGGCGCGTTCACAAACCACTGCAGGGTGATAACCGTGGTCAGAA 143
|||
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
|||
Db 144 GAAGTAGCCCTTTATGAATCATTTCTTCAACAATATATTCAGAACACACATACGC--- 200
|||
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
|||
Db 201 AAATTCCTTCATATATTATGACCAACAAATCATG-----AAGGCATCCACTGGCTCT 254
|||
QY 84 ProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaProCysValAla 103
|||
Db 255 GACCATCTCATCATGGTGTGCAAGATCTTACATCAGCTCATGTCAACCCATCTGTGAATG 314
|||
QY 104 AspIleLysIleGlyAlaIleThrProProSerSerProGluProTyrIleAlaLys 123
|||
Db 315 GACATCAAAATCGGTCAGAACATGGCGCCGAGAGCTTCCGAGCGGTACATTGCAAA 374
|||
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
|||
Db 375 TGCTTAAAAAGGATAGGGAAGCACAGTATTCCTATGGGATTCAGGATCTCCGGGCTG 434
|||
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
|||
Db 435 CAAAGTCTATATCGATGGGTGAGGCTTTTATAAGCCTCATAGAAATTTACATGCGTAA 494
|||
QY 163 MetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAlaAspGlu 182
|||
Db 495 ACCGGCCCGAGCTGATGTAGACTACTCTTAGAAATTTGTTCTTAAACCCGCTCGCA 554
|||
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
|||
Db 555 GAGATGGAATCGGCACAGGCCCTAGCCCGGATGTTCTTTAGCATCTTTTGTATTGCT 614
|||
QY 195 GlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThr 214
|||
Db 615 GGGCCCTAATGGGATATTAGCTCAACTGATGGAATTTGAAGACATGGTTTGAAGATCAAA 674
|||
QY 215 LeuPheHisPheThrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAla 234
|||
Db 675 ATTTACCACTTCTATGCTGTTGTTTTCATCTTTTGTTCATCTTTTGAAGAGGTGGTGTAA 734
|||
QY 235 GlyGlyAspGlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGly 254
|||
```



```
Db 735 GGTGCT---CGGTCAACACGAGAGTCAAACTTATTGATTGCTCATGTTACAGATGGT 791
Qy 255 AspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSer 274
Db 792 AATGGGTGTTATTGATCAAAATTTCTTGGGTGGGCTCTGTTCTTTGATAAAGTTTCATTTCT 851
Qy 275 AspIleValProGluThr 280
Db 852 GACATACTTTCGGGAGACA 869

RESULT 12
US-10-042-894A-11
; Sequence 11, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1195
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)...(1048)
US-10-042-894A-11

Alignment Scores:
Pred. No.: 2.44e-76 Length: 1195
Score: 722.50 Matches: 149
Percent Similarity: 66.31% Conservative: 38
Best Local Similarity: 52.84% Mismatches: 84
Query Match: 47.50% Indels: 11
DB: 14 Gaps: 6

US-10-042-894A-8 (1-289) x US-10-042-894A-11 (1-1195)
Qy 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 119 CTCAAGTCCGGATCATCAAGTCGCCGGTCCCGGGGACGGGGGAAAGCTGGGGCCA 178
Qy 24 LeuIleAspGlySerGlyLeuPheTyLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 179 CTGCTGGATGATTCGGGCCCGCTCTAAGCCTCTCCAGAGCGATCATCGCGGAGACAG 238
Qy 44 GluValAlaPheTyrgluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 239 GAAGTGGCCCTTTACGAGTCACTTATTTCAATACCGAGATCCCGAGTCACATTCGC--- 295
Qy 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 296 AATTTCTTTCTGGTTTCACGGAACTAAGACTATT-----GAGGGCTCTGATGGATCG 349
Qy 84 ---ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db 350 GGTCCTCAACCTCACTGGTTCTCGAGGATCTCGTCTCGGTGCGACCAACCCATCTCTC 409
Qy 103 AlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyrlleAla 122
Db 410 ATGGACATCAAGACTGGATCCAGAACATGTTATCGGAGGCGCTCTGAGGAGTACATCAA 469
Qy 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
```

```
Db 470 AAGTGTCTAGAGAAAGATCGAAATACCAAGCGTTTCATTGGGTTTTAGGATTTCTGGG 529
Qy 143 ValArgValVal---GlyProGluGlyAlaValTTPArgThrGluArgProGluValLys 161
Db 530 CTAAAGGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAGAAGGTTGTTTAT 589
Qy 162 AlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAlaAsp 181
Db 590 AGCTTTAATCGGACCGGTGTCAAGTCGGCTCTGAGGAAGTTTGTCTTCTTCCAACTTGCT 649
Qy 182 GluGly-----MetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 197
Db 650 CTGGGTCCAAATGTGGATCCGATTGTTGTATGTCATCAAAAGTTTACTGTCCCGGGGT 709
Qy 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGluGlnThrLeuPheHis 217
Db 710 GGAATTTTGGCAAAATGCTTTCAGCTGAAGGAATGTTTGGGTTTCAGAGCAATATATCAC 769
Qy 218 PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAsp 237
Db 770 TTCTATTCTTGTTCATCTATTATATGACAGGGAGTCTGCTTTGGACGGC----- 823
Qy 238 GlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 824 TGTGCACACCCGAAAGTTAACTGGTGGACTTTGCACATGTGATGGATGCCACGGCGTG 883
Qy 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 884 ATCGATCACAACTTCTTGGGTGGCCTCTGTCTGTAAATCAAGTTTATACGTGACATGCT 943
Qy 278 ProGlu 279
Db 944 GATGAA 949

RESULT 13
US-10-042-894A-9
; Sequence 9, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)...(851)
US-10-042-894A-9

Alignment Scores:
Pred. No.: 5.01e-67 Length: 1105
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservative: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 14 Gaps: 8

US-10-042-894A-8 (1-289) x US-10-042-894A-9 (1-1105)
```

```
QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAsp 26
Db 24 CCGGAGCACCAGTGGCGGGCACAGGCAAGGACCGAATCTCTGGGCCCACTCTCGTCGAC 83
QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42
Db 84 GATTTTGGAAAAATTTACAAGCCCTCCAGACCAAAAGACGACGACACCGCGGCTCC 143
QY 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArg 62
Db 144 ACCGAATCTCTCTTTTACACCTCTCGCGGCC--GCGGCCACGACTACTCCATCCGC 200
QY 63 AspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly 82
Db 201 ---TCCTTCTTCCCGCCCTTTCAGGACCCCGCTCTCGACGCCCTCCGACGGCTCCGCT 257
QY 83 GluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysVal 102
Db 258 ---CCGACCCCTCACCTGGTCTCTGGAGGACCTCTCTGGCGGTACTCTCAAACCCCTCCGTC 314
QY 103 AlaAspIleLysIleGlyAlaIleThrTTPProSerSerProGluProTyrIleAla 122
Db 315 ATGGAGCTAAAGATCGGCTCCAGAACCTGGCACCTGGGAGACTCCGAGGACTACATCTGC 374
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 375 AAGTGCCTGAAGAGGACAGAGAGTCTCTAGCTTGCCTTGGTTTCAGAAATCTCGGGA 434
QY 143 ValArgValValGlyProGluGlyAlaValTTPArgThrGluArgProGluValLysAla 162
Db 435 GTCAAG-----GACTCTATCTCTCTCTGGGAACCTACCAGGAAATCTCTCCAGTGT 485
QY 163 MetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer----- 178
Db 486 CTATCGCCCATGGTGTGACCTGTCTTCAACAAGTTCTGTTTCTCTTAATATCAAC 545
QY 179 ValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 198
Db 546 CATGATGATCATCATCCGATTCGGCTTCGCACACGAGGAGTCTAC-----GCGGCC 596
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGluGlnThrLeuPheHisPhe 218
Db 597 GTTTTGGAGCGCTTGCAAGAGCTCAAGGACTGGTTCGAGGTTTCAGACGGTGTATCACTTC 656
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGly 238
Db 657 TATTTCTGTCTGTTCTTGTGGTACGAGAGGAT-----CTAGGGAAGGGAAGCT 710
QY 239 GlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIle 258
Db 711 ACCAACCCCTCTGTCAAACTCGTTGACTTTGCACACGTGTGTGACGAGGAACGGTGTCA 770
QY 259 AspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 771 GATCAACAACCTCTTGGGTGGCCCTTGTCTTCATCAAGTTCTCTCAAGGATATCTCTA 827

RESULT 14
US-10-042-599-106467
; Sequence 106467, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 106467
; LENGTH: 1345
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67156C.1
; US-10-424-599-106467
```

```
Alignment Scores:
Pred. No.: 6,47e-67 Length: 1345
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservatives: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 18 Gaps: 8
```

US-10-042-894A-8 (1-289) x US-10-424-599-106467 (1-1345)

```
QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAsp 26
Db 229 CCGGAGCACCAGTGGCGGGCACAGGCAAGGACCGAATCTCTGGGCCCACTCTCGTCGAC 288
QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42
Db 289 GATTTTGGAAAAATTTACAAGCCCTCCAGACCAAAAGACGACGACACCGCGGCTCC 348
QY 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArg 62
Db 349 ACCGAACTCTCTTTTACACCTCTCTCGCGGC--GCGGCCACGACTACTCTCCATCCGC 405
QY 63 AspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly 82
Db 406 ---TCCTTCTTCCCGCCCTTTCAGGACCCCGCTCTCTGGAGCCCTCTCGACGGCTCCGCT 462
QY 83 GluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysVal 102
Db 463 ---CCCCACCCCTCACCTGTCTCTGGAGGACCTCTCTCGGGTACTCTCAAAACCCCTCCGTC 519
QY 103 AlaAspIleLysIleGlyAlaIleThrTTPProSerSerProGluProTyrIleAla 122
Db 520 ATGGAGCTAAAGATCGGCTCCAGAACCTGGCACCTGGGAGACTCCGAGGACTACATCTGC 579
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 580 AAGTGCCTGAAGAGGACAGAGAGTCTCTAGCTTGGCTTGGGTTTCAGAAATCTCGGGA 639
QY 143 ValArgValValGlyProGluGlyAlaValTTPArgThrGluArgProGluValLysAla 162
Db 640 GTCAAG-----GACTCTATCTCTCTCTGGGAACCTACCAGGAAATCTCTCCAGTGT 690
QY 163 MetAspThrAlaGlyValArgValLeuArgValTyrValSerSer----- 178
Db 691 CTATCGCCCATGGTGTGACCTGTCTTCAACAAGTTCTGTTTCTCTTAATATCAAC 750
QY 179 ValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 198
Db 751 CATGATGATCATCATCCGATTCGCTTTCGCAACGAGGAGTCTAC-----GCGGCC 801
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGluGlnThrLeuPheHisPhe 218
Db 802 GTTTTGGAGCGCTTGCAAGAGCTCAAGGACTGGTTCGAGGTTTCAGACGGTGTACTACTTC 861
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGly 238
Db 862 TATTTCTGTCTGTTCTTGTGTGTCAGGAGGAT-----CTAGGGAAGGGAAGCT 915
QY 239 GlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIle 258
Db 916 ACCAACCCCTCTCGTCAAACTCGTTGACTTTGCACACGTGTGTGAGCGGAACGGTGTCA 975
QY 259 AspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 976 GATCAACAACCTCTTGGGTGGCCCTTTGTCTTCATCAAGTTCTCTCAAGGATATCTCTA 1032

RESULT 15
```

US-10-425-115-74436  
; Sequence 74436, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 74436  
; LENGTH: 776  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(776)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_167884C.1  
US-10-425-115-74436

Alignment Scores:  
Pred. No.: 5,73e-63 Length: 776  
Score: 609.00 Matches: 117  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.15% Mismatches: 0  
Query Match: 40.04% Indels: 0  
DB: 20 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-425-115-74436 (1-776)

QY	172	LeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla	191
DB	76	GTCCGGCGCTACGTGTCNTCCGTTGCCGACGAGGGGATGGACTGTGGCTCGCGCGGCG	135
QY	192	ValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaIrrPheGlu	211
DB	136	GTGTACGGAGAAAGGTGGAGTCTTGTACAGCTCGCGAGCTCAAGGCGTGGTTCGAG	195
QY	212	GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAla	231
DB	196	GACGAGACTCTGTCCACTTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGTCA	255
QY	232	ValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisVal	251
DB	256	GTCCGACGAGCGGAGATGGGGGTGGGGTGACGCGTGAAGCTGGTGGACTTTGCCCATGTG	315
QY	252	AlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyLeuCysSerLeuIleLys	271
DB	316	GCCGAGGGTGATGGGGTGATTGACCACAACTTCTCGGCGGGCTCTGCTCGCTGATCAAG	375
QY	272	PheValSerAspIleValProGluThrProHisThrGlnProLeuGlyProSer	289
DB	376	TTGCTTTCTGACATTGTTCCGAGACTCTCTCATACGAGCCTTTGGGTCTCTTCT	429

Search completed: June 17, 2005, 09:05:18  
Job time : 2137 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 03:46:13 ; Search time 205 Seconds  
(without alignments)

2306.751 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVACHRASASK.....IKFVSDIVPETPTQLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111951\_8314/app\_query.fasta\_1.455  
-DB=Issued Patents NA -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894 @CGN\_1\_69 @runat\_15062005\_111951\_8314 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	12.7	1341	4	US-09-023-655-809
2	189	12.4	1748	4	US-09-620-312D-712
3	182	12.0	1863	4	US-09-774-528-103
4	178.5	11.7	1893	4	US-09-774-528-102
5	145.5	9.6	1781	4	US-09-949-016-4634
6	139	9.1	963	4	US-09-248-796A-5436
7	113.5	7.5	1899	4	US-09-252-991A-6525
8	112.5	7.4	4496	4	US-09-949-016-2348
9	112.5	7.4	4505	4	US-09-949-016-323
10	108	7.1	14555	4	US-09-902-540-1096
11	107.5	7.1	954	4	US-09-248-796A-4664
12	107.5	7.1	984	4	US-09-252-991A-12872

C 13	107.5	7.1	4078	4	US-09-016-434-1132	Sequence 1132, Ap
C 14	106.5	7.0	6000	1	US-08-348-006B-6	Sequence 6, Appli
C 15	106.5	7.0	6000	2	US-08-800-825A-6	Sequence 6, Appli
C 16	106.5	7.0	6000	3	US-09-158-657-6	Sequence 6, Appli
C 17	106.5	7.0	6000	5	PCT-US94-10166-6	Sequence 6, Appli
C 18	106.5	7.0	411529	3	US-09-103-840A-1	Sequence 1, Appli
C 19	105.5	6.9	36412	4	US-08-311-731A-132	Sequence 132, App
C 20	104	6.8	1917	2	US-08-637-899-2	Sequence 2, Appli
C 21	103.5	6.8	1896	4	US-09-252-991A-11331	Sequence 11331, A
C 22	101.5	6.7	8332	4	US-09-902-540-927	Sequence 927, App
C 23	101	6.6	1161	4	US-09-902-540-7636	Sequence 7636, Ap
C 24	101	6.6	2460	4	US-09-902-540-4178	Sequence 4178, Ap
C 25	101	6.6	12419	4	US-09-902-540-1059	Sequence 1059, Ap
C 26	101	6.6	22301	4	US-09-902-540-1208	Sequence 1208, Ap
C 27	100.5	6.6	960	4	US-09-252-991A-12565	Sequence 12565, A
C 28	100.5	6.6	2514	3	US-09-144-914-3	Sequence 3, Appli
C 29	100.5	6.6	2568	4	US-09-949-016-1555	Sequence 1555, Ap
C 30	100.5	6.6	3446	4	US-09-620-312D-653	Sequence 653, App
C 31	100.5	6.6	17173	4	US-09-902-540-1122	Sequence 1122, Ap
C 32	100.5	6.6	43507	4	US-09-949-016-13297	Sequence 13297, A
C 33	100	6.6	1425	4	US-09-489-039A-1509	Sequence 1509, Ap
C 34	100	6.6	1491	4	US-09-252-991A-6021	Sequence 6021, Ap
C 35	100	6.6	2016	4	US-09-489-039A-1492	Sequence 1492, Ap
C 36	99.5	6.6	5663	4	US-09-902-540-839	Sequence 839, App
C 37	99.5	6.5	1681	4	US-09-270-767-1662	Sequence 1662, Ap
C 38	99.5	6.5	1681	4	US-09-270-767-16944	Sequence 16944, A
C 39	99	6.5	1470	4	US-09-489-039A-6625	Sequence 6625, Ap
C 40	99	6.5	19068	4	US-09-902-540-1123	Sequence 1123, Ap
C 41	99	6.5	23856	4	US-09-949-016-15732	Sequence 15732, A
C 42	98.5	6.5	7201	4	US-09-902-540-914	Sequence 914, App
C 43	98	6.4	12194	4	US-09-902-540-1091	Sequence 1091, Ap
C 44	98	6.4	23677	4	US-09-902-540-1218	Sequence 1218, Ap
C 45	98	6.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-023-655-809  
; Sequence 809, Application US/09023655  
; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS: 1508

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:



```
QY 97 eGlnAlaProCysValAlaAspIleValIleGlyAlaIleThrTrpProProSerSerPr 117
Db 825 TGAGGTGCGTGTGTCTTCAAGATGGGACACGACACATGATGATGCTTC 884
QY 117 oGluProTyrlleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuG1 137
Db 885 AGAGGAGAGCGCCCAACAGATCCGANAATGCACGAGCACATCTCAGTCATGCG 944
QY 137 yPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluAr 157
Db 945 TGTGCGTGTGTGTGCGATGAGGTGTAC-----CAAGCAGGCGAG-TGGCGAGTCATGT 997
QY 157 gProGluValIleAlaMetAspThrAlaGlyValArgArgValLeuArgGlyTrpValSe 177
Db 998 TCATGAACA-AGTACCATTGACGAGGAGTATCGGTGAGGCGTTCAAGGAGGACATTTTC 1056
QY 177 rSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyLysG1 197
Db 1057 CAGTTCTTCCACATGGCGGTACTGCGCGTG-----AACTCCTG-GG 1100
QY 197 yGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHi 217
Db 1101 CCCTGTGCTCAAGAAGCTGACTGAGCTCAAGCGAGTGTTCGAGCGACAGGAGTCTTACC 1160
QY 217 sPheTySerAlaSerIleLeuLeuGlyTyArgP-----228
Db 1161 CTTCTACTCAAGTCCCTGCTGTGTTATTTATGATGGCAAGGAGCGGCCGAACTGGTCT 1220
QY 229 -----AlaAlaAlaValAlaAlaG1 235
Db 1221 GGACTCAGATGCTGAGGATTTGGAGGACCTGTACAGGATCAGCTGATGATGCTGCTGG 1280
QY 235 yGly-----AspGlyGlyValThrValLysLeuValAspPheAl 249
Db 1281 TGCCTATGCTTACAAACCCATCGCGCGAGCTCTGTAGATGTGGCATGATGCACTTTC 1340
QY 249 aHis-----ValaGluGlyAspGlyValI1 258
Db 1341 ACACACACCTGCAGGCTGTATGCGGAGGACACCGTGTGATGAGGCGGCGAG----- 1392
QY 258 eAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValPr 278
Db 1393 -GATGCTGGCTATATCTTCGGGCTCCAGAGCCTGATAGACATTTCTACAGAGTAAAGTGA 1451
QY 278 oGluThr 280
Db 1452 GGAGAGT 1458
```

## RESULT 3

```
US-09-774-528-103
; Sequence 103, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
```

```
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
```

```
; SEQ ID NO 103
; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (367)..(1593)
US-09-774-528-103
```

```
Alignment Scores:
Pred. No.: 7,4e-11 Length: 1863
Score: 182.00 Matches: 67
Percent Similarity: 44.53% Conservative: 47
Best Local Similarity: 26.17% Mismatches: 82
Query Match: 11.97% Indels: 60
DB: 4 Gaps: 10
```

```
US-10-042-894A-8 (1-289) x US-09-774-528-103 (1-1863)
```

```
QY 62 ArgAspThrPhePhePro-----ArgPheHisGlyThrArgLeuLeu 75
Db 868 AGGAAGAGCTTCAACCCGTGGGCGCTTGCATATGCCACGAGCCACCTGACCCGCTGTC 927
QY 76 ProThrGluAlaGlnProGlyGluProHisProHisValLeuValLeuAspLeuLeuAla 95
Db 928 -----TCCGATGATCCAGAGAACAAAGCGGTCTTGTGTGGAAATGATGTGTC 978
QY 96 GlyPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProProSer 115
Db 979 CAGTACACGATCCCTGTGTCTGATCTGAAGATGGGACCGGAGCGGAGCGATGAT 1038
QY 116 SerProGluProTyrlleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeu 135
Db 1039 GCATCGAGGAGAGAGGCGCCCATGAGGAAGTGTGCGCAGAGCACCTCAGCCTGC 1098
QY 136 LeuGlyPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpArgThr 155
Db 1099 CTGGGTGTGGCATCTGCGGCGATGCGAG-----GTTTATCAACA 1137
QY 156 GluArgProGluValLysAlaMetAspThrAlaGlyValArgValLeuArgGlyTyr 175
Db 1138 GATAAGAAGTACTTCTCTGCAAGAC-----AAGTACTAT 1173
QY 176 ValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaValTrpGlyGly 195
Db 1174 GGAAGAAAACTCTCAGTGGAGGGGTTTCAGACAAAGCCCTCTATCAGTTCATACATGGA 1233
QY 196 Lys-----GlyGlyValLeuSerGlnLeuArgGluLeuLeuAla 208
Db 1234 AGCCACTCGGAGGAGCTCTCGGAGCCCATCTGCACAGCTCCGGGCGCTCTCTCT 1293
QY 209 TrpPheGluGluGlnThrLeuPheHisPheTySerAlaSerIleLeuLeuGlyTyrAsp 228
Db 1294 GTCATTAGGAGCCAGATTTCATACCGCTTCTATTCAGCTCTCTCTCTTGTCTATGAT 1353
QY 229 -----AlaAlaAlaVal 232
Db 1354 GGGCAGGAACCCAGAAAGAGCCCGGAGCGCCGATCTCTCAGGAGTCCCGAGGCA 1413
QY 233 AlaAlaGlyGlyAspGlyGlyValThr-----ValLysLeuValAspPheAla 249
Db 1414 GCCCAGGTAGCTCTCCCGGTGTCTCACCAGGTTGACATCCGATGATGATTTGCT 1473
QY 250 HisValAla-----GluGly-----AspGlyValIleAspHis 260
Db 1474 CATACCATACAGGGGCTACTGGAATGAGCACACACCATCTACGATGACCA---GACCT 1530
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276
Db 1531 GGCTATATTTTGGCCTGGAAACCTCATCAGGATCTCTGAGGATATC 1578
```





Db 752 CTCAGGACCTGCTCGATGCTTCGAGGACCTTGTGTGCTGCGACTGCAAAATGGCGTC 811  
QY 110 IleThrTrp----- 112  
Db 812 AGGACTTACCTAGAGGAGGAGTGACCAAGGCCGCTGAGCGGCCCAAGCTGCGAGAGGAC 871  
QY 113 -----ProProSerSerPro-----GluProTyrIleAla 122  
Db 872 ATGTACAAGAAATGCTGGCGGTGGATCTCTGAAGCTCCACGAGGAGGACGCGCGAG 931  
QY 123 LysCysLeuAlaLysAspArgGly-----ThrThrSerValLeu 135  
Db 932 CGGCGCTACCAAGCGCGCTACATGCTAGTGGCGGAGGAGGATCAGTCCAGCACCCACC 991  
QY 136 LeuGlyPheArgValSerGlyValArgValVal-----GlyProGluGlyAlaValTrpArg 154  
Db 992 CTCGGCTTCGCGATCGAGGCGCATCAAGAAAGCGAGCGGCTCTCGACGACCCGATTCNAG 1051  
QY 155 ThrGluArgProGluValLysAlaMetAsp-----ThrAlaGlyValArg 169  
Db 1052 ACTACCGAAGCCGAGAGCGAGGTCTTCGGCTCTTTGAAGAGTTTGTGCAGAGGATGAG 1111  
QY 170 ArgValLeuArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAla 189  
Db 1112 GAAGTCTGAGCGGTAT----- 1129  
QY 190 AlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaTrp 209  
Db 1130 -----CTGAACCGCTCGAGCAGATCCGGGACACC 1159  
QY 210 PheGluGlnThrLeuPheHisPheTyr-----SerAlaSerIleLeuGly 226  
Db 1160 CTGAGGATCCGAGTCTTCAGAGCAGCAGGTGATCGCAGCTCGCTCTTTGTG 1219  
QY 227 TyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuVal 246  
Db 1220 CAGATCAGTCCATCGCGCGCG-----GTGTGCTCATC 1255  
QY 247 AspPheAlaHisValAlaGlu-----GlyAspGly-----ValIleAspHis 260  
Db 1256 GACTTCGGCAAGACCAGCGCCCTCCCGATGGCCAGATCCTGGACCACCGCGGCCCTGG 1315  
QY 261 -----AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal 273  
Db 1316 GAGAGGCGCAACCGCAGGACGCGCTATTGTGGCGCTGGACAACTCTCATTTGGCATCCTG 1375  
QY 274 SerAspIleValProGlu-----ThrProHisThrGlnProLeuGlyProSer 289  
Db 1376 GCCAGCTGCTGCTGAGAGATGAGCTGGACTCCT-----GTCCCGCGCGCGCTCA 1426

## RESULT 6

US-09-248-796A-5436  
; Sequence 5436, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE OF INVENTION: 107196.132  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 5436  
; LENGTH: 963  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-5436  
Alignment Scores:

Pred. No.: 2,59e-06 Length: 963  
Score: 139.00 Matches: 67  
Percent Similarity: 37.83% Conservative: 48  
Best Local Similarity: 22.04% Mismatches: 105  
Query Match: 9.14% Indels: 84  
DB: 4 Gaps: 12  
US-10-042-894A-8 (1-289) x US-09-248-796A-5436 (1-963)  
QY 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAla-----SerAlaSer 19  
Db 223 TCCCAATGACCCGCTTTAAACAAGTGGGAGGTACATACACCCATATTTCTCAATATCA 282  
QY 20 Lys-----LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGln 36  
Db 283 AAGCAAGAAATATGCAAGAGCTCAACATACCGAAGAAATTTCTCAAGAAATC--- 339  
QY 37 AlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAla 56  
Db 340 -----TCCAAGAATCATCTTTATTTAT----- 363  
QY 57 ValProAlaArgIleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuPro 76  
Db 364 -----TATATGCCAAGGTATAAAGGTAGTAAT----- 390  
QY 77 ThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAspLeuAlaGly 96  
Db 391 -----GGATCCCAANTATTTCTAGAGATCTTACATCTCAA 426  
QY 97 PheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSer 116  
Db 427 ATGAGAACTCTTGATATTTGATTTAAAGATGGGCAAGACATATATGGGTGTAATGCA 486  
QY 117 ProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeu 136  
Db 487 ACTATCACCACAAACAATCACATCAGCTAGCTAAAGCCAGACGACAAACAAGAAATG 546  
QY 137 GlyPheArgValSerGlyValArgValValGlyProGlyGlyAlaValTrpArgThrGlu 156  
Db 547 GGTGTTAGAAATATGCGACTTCAAAATCTTTAATATCAAAACAATATTTCTATCAGAT 606  
QY 157 ArgProGluValLysAlaMetAspThrAlaGlyValArgValLeuArgTyrVal 176  
Db 607 -----AAATATCTTGAAGAAAGATA 627  
QY 177 SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLys 196  
Db 628 ACTGTTGGTAAACAGTTTGGCAAA-----ATTTTAGCAAAATTTTATATAATGGCCAT 681  
QY 197 Gly-----GlyValLeuSerGlnLeuArgGluLeuLys 207  
Db 682 GACATTTATTCATTATTAATCGTATCCCGCATTTAATCGATCACTTAAGAATATATAC 741  
QY 208 AlaTrpPheGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyr 227  
Db 742 ACAATATTACTGGATTACCGAGGTATAGATGTATGGTAGTCTGATATATTATTAATGAT 801  
QY 228 AspAlaAlaAlaValAlaAlaGlyGlyAspGly-----GlyGlyValThrValLys 244  
Db 802 GAA-----GGAGGTGAAGATAAATCTGGAAATCAAGTAAAGTGAAG 843  
QY 245 LeuValAspPheAlaHis-----ValAlaGluGlyAspGlyValIleAspHisAsnPheLeu 263  
Db 844 ATTATAGATTTTGTCTAATCAGTATAGCAGCGCCGACACATATATAGTAAT----- 894  
QY 264 GlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThrProHisThr 283  
Db 895 -----GTGACTGTTGCACCACCAACCCCTGATTCC 924  
QY 284 GlnProLeuGly 287  
Db 925 CCCAAACTGGGG 936

## RESULT 7

US-09-252-991A-6525/c  
; Sequence 6525, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 6525  
; LENGTH: 1899  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6525

Alignment Scores:  
Pred. No.: 0.00663 Length: 1899  
Score: 113.50 Matches: 93  
Percent Similarity: 33.51% Conservative: 35  
Best Local Similarity: 24.35% Mismatches: 115  
Query Match: 7.46% Indels: 139  
DB: 4 Gaps: 20

US-10-042-894A-8 (1-289) x US-09-252-991A-6525 (1-1899)

QY 12 AlaGlyHisArgAlaSerAlaSerLeuGlyProLeuIleAspGlySerGlyLeuPhe 31  
DB 1515 GCGCGCACCGCGTACTTGGTGGCGGTGTAGGCGCT-----CATCGCGGTGGGC 1465  
QY 32 TyrLysProLeuGlnAlaGlyAsp-----ArgGlyGluHisGluVal 45  
DB 1464 GAACAGCCCGAGACCGCTGGAGACATTGACACGTGGCGCTTGGCGAGCGCTTCAGGTG 1405  
QY 46 AlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAspThrPhe 65  
DB 1404 CGAAGGAACGCGTGTGGTCCGCTTCCACCGCCAGAGAGTTGATGTTTCATGATCCACTC 1345  
QY 66 PheProArg-----PheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81  
DB 1344 GTATTCGAGTAGTCGCTGCTTCCACCGTACCGCGTGGCGACGCCCGCGGTGTGTGAA 1285  
QY 82 GlyGluProHisPro-----HisLeuVal 89  
DB 1284 AATCAGTTTACCCGCGCATGTTGCGTGGCGCTTGTGGCCAGCGCTGCACCTGCTC 1225  
QY 90 -----LeuAspAsp-----LeuLeu 94  
DB 1224 GCGGTGCGCACCGCTGCACGACCGCGGTGGAGACGCTACCGCTGCTGGCGAGCAATTG 1165  
QY 95 AlaGlyPheGlnAlaProCysValAlaAspIleLysLeuGlyAlaIleThrTrpProPro 114  
DB 1164 GCGGGTTTCTCGAGAGCTGCGGCATGACGTGCGCCAGCGCAGGTGACAAACCTGCT 1105  
QY 115 Ser--SerProGluProTyrIleAlaLysCysLeuAlaLysAspArg----- 129  
DB 1104 GCCAGCTCCAGCGCAGGCGGACCGATGCCGAGCTGCGCGGTGATGGCGCGAC 1045  
QY 130 -----GlyThrThrSer-----ValLeuLeuGlyPheArgVals 141  
DB 1044 TTTGTTTTCCAGCAGCTTATCGGGGTACCCCTTCGGCCTGTGGCAGGCTTCGTTGCTC 985  
QY 141 exCly-----ValArgValValGly----- 147  
DB 984 AGGCGCGGAGCGGTGCTGTTGAGTGTGTGGCGGAGCGGTGCAACTGCCGGGTGAG 925  
QY 148 -----ProGluGlyAlaValTrpArgThrGluArgProG 159

DB 924 GCACGGAAGCGCCAGGTGAATCCCGCCACAGGGGTGAGTTGCGGCGCTGACCGGATGC 865  
QY 159 luValLysAlaMetAspThr-----AlaGlyValArgArgValLeuArgA 174  
DB 864 AGGTACCAAGCTTGTGCAACCGCCCGCTCCAGACGCTGCGGTCCAGCTTGCCTGGAGG 805  
QY 174 tGTYrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAla-----AlaAlav 192  
DB 804 TATTCGTTGTAGCGTTCCTGCACCGCGGCTTGACCTCGAGACTGAGCAGTTGCGGCGC 745  
QY 192 alTyrGly----- 194  
DB 744 TTCATCAGCTTCAGCGCATCGAGGAGTAGGCGATCTGCGACTCGATCATGTAGACCATC 685  
QY 195 -----GlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaIlePheGluG 212  
DB 684 GAGTTGTGCGCGCAGGCGGTGTTGCGTCCCAT-----CAGG 649  
QY 212 luGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAla----- 229  
DB 648 AAGAACAGTTGGGAAGCGCGGTGTGTAGTGCCTTGTAGGCTTCCGGGCGCTTGGTGC 589  
QY 230 -----AlaAlaValAla-----AlaGlyGlyAsp----- 237  
DB 588 CAGCTGTCCAGCAGGTGCGGCGCGTCCGACCGAAGACCACTCCGCGGGGATAGGATCG 529  
QY 238 --GlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAla----- 252  
DB 528 TTGGCGGTGAAGCCGCTACCGAAGATGATCGCGTCCGACTCCGCTTCCCGCGGTTGCCG 469  
QY 253 -----GluGlyAspGlyValIle-----AspHisAsnPheLeuGlyG 265  
DB 468 TCGACGATTCATTGCGCGTGACGCGCGGATGCTTCCGTGATCACCGTGGAGTTGGCG 409  
QY 265 ly 265  
DB 408 GC 407

## RESULT 8

US-09-949-016-2348  
; Sequence 2348, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2348  
; LENGTH: 4496  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-2348

Alignment Scores:  
Pred. No.: 0.0325 Length: 4496  
Score: 112.50 Matches: 59  
Percent Similarity: 34.98% Conservative: 40  
Best Local Similarity: 20.85% Mismatches: 87  
Query Match: 7.40% Indels: 97  
DB: 4 Gaps: 15

US-10-042-894A-8 (1-289) x US-09-949-016-2348 (1-4496)

```
QY 65 PheProArgPheHisGlyThrArgLeuProThrGluAlaGlnProGlyGluPro 84
Db 807 TTCTGTAACCTCCCTACCATGGG-----GATGTGTGAGGACGGGGAGCGC 851
QY 85 HisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAsp 104
Db 852 TACAACACAG-----ATGGACGACCTGCTGGCCGACTTCGACTCGCCCTGTGTGATGGAC 905
QY 105 IleValIleGlyAlaIleThrTrp-----
Db 906 TGCAAGATGGGAATCAGGACCTACTCTGGAGGAGGCTCACGAAGGCCCGGAAGAAGCC 965
QY 113 -----ProProSerPro----- 117
Db 966 AGCTCGGAGGACATGTACCAAGAGATGATCAGGTGGACCCCGAGGCCCCACCGGAG 1025
QY 118 GluProTyrIleAlaLysCysLeuAlaLysAspArgGly----- 130
Db 1026 GAGGAAAAAGCACAGCGGGCTGTGACCAAGCCAGGTACATGCGTGGCGGGAGACCATC 1085
QY 131 ThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValGlyProGluGly 150
Db 1086 AGCTCCAGCCGCCCTCTGGGGTTTCAGGATCGAGGAATCAAG-----AAAGAAGACGGC 1139
QY 151 AlaValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArg 170
Db 1140 ACCGTG-----AACCGGAGCTTCAGAGACCAACAAAGAGGAGGACGATCAGCGAG 1190
QY 171 ValLeuArgArgTyrValSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 190
Db 1191 GCCTTCAGAGAGTTCACTAAA----- 1211
QY 191 AlaValTyrGlyGlyLysGlyGlyValLeu-----SerGlnLeuArgGluLeu 206
Db 1212 -----GGAACCAATACATCTGATCGCCCTATCGGACCGGCTGAAGGCCATT 1259
QY 207 LysAlaTrpPheGluGlnThrLeuPheHisPheTyr-----SerAlaSerIle 223
Db 1260 CGAACCACTCTAGAAGTTTCTCCCTCTTCAAGTGCACGAGGTATGTGGCAGCTCCCTC 1319
QY 224 LeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrVal 243
Db 1320 CTCTTCATCCACGACAGAGGAAACAGGCC-----AAAGTG 1355
QY 244 LysLeuValAspPheAlaHis-----ValAlaGluGly----- 254
Db 1356 TGGATGATGACTTTGGAAAAACACCGCCCTGCTGAGGCCAGACCTGCGAGCATGAC 1415
QY 255 -----AspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIle 270
Db 1416 GTCCCTGGCAGAGGGGAACCGGGAGGATGGTACCTCTCGGGGCTCAATACCTCGTC 1475
QY 271 LysPheValSerAspIleValProGluThrPro-----HisThrGlnPro 285
Db 1476 GACATCTGACCGAGATGTCACGAGATGCCCTGCTGCTGAGCTGCCACGCCCCCT 1535
QY 286 LeuGlyPro 288
Db 1536 ---GGCCCC 1541
```

## RESULT 9

```
US-09-949-016-323
; Sequence 323, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323
; LENGTH: 4505
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-323
```

```
Alignment Scores:
Pred. No.: 0.0326 Length: 4505
Score: 112.50 Matches: 59
Percent Similarity: 34.98% Conservative: 40
Best Local Similarity: 20.85% Mismatches: 87
Query Match: 7.40% Indels: 97
DB: 4 Gaps: 15
```

US-10-042-894A-8 (1-289) x US-09-949-016-323 (1-4505)

```
QY 65 PheProArgPheHisGlyThrArgLeuProThrGluAlaGlnProGlyGluPro 84
Db 810 TTCTGTAACCTCCCTACCATGGG-----GATGTGTGAGGACGGGGAGCGC 854
QY 85 HisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAsp 104
Db 855 TACAACACAG-----ATGGACGACCTGTGGCCGACTTCGACTCGCCCTGTGTGATGGAC 908
QY 105 IleValIleGlyAlaIleThrTrp----- 112
Db 909 TGCAGATGGGAATCAGGACCTACTCTGGAGGAGGAGCTCACGAAGGCCCGGAAGAAGCC 968
QY 113 -----ProProSerPro----- 117
Db 969 AGCTCGGAGGACATGTACCAAGAGATGATCGAGGTGGACCCCGAGGCCCCACCGAG 1028
QY 118 GluProTyrIleAlaLysCysLeuAlaLysAspArgGly----- 130
Db 1029 GAGGAAAAAGCACAGCGGGCTGTGACCAAGCCAGCGGTACATGCGTGGCGGGAGACCATC 1088
QY 131 ThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValGlyProGluGly 150
Db 1089 AGCTCCAGCCGCCCTCTGGGGTTTCAGGATCGAGGAATCAAG-----AAAGAAGACGGC 1142
QY 151 AlaValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArg 170
Db 1143 ACCGTG-----AACCGGAGCTTCAAGAGACCAACCAAGAGGAGGAGGTACCGGAG 1193
QY 171 ValLeuArgArgTyrValSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 190
Db 1194 GCCTTCAGAGAGTTCACTAAA----- 1214
QY 191 AlaValTyrGlyGlyLysGlyGlyValLeu-----SerGlnLeuArgGluLeu 206
Db 1215 -----GGAACCAATACATCTGATCGCCCTATCGGACCGGCTGAAGGCCATT 1262
QY 207 LysAlaTrpPheGluGlnThrLeuPheHisPheTyr-----SerAlaSerIle 223
Db 1263 CGAACCACTCTAGAAGTTTCTCCCTCTTCAAGTGCACGAGGTATGTGGCAGCTCCCTC 1322
QY 224 LeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrVal 243
Db 1323 CTCTTCATCCACGACAGAGGAAACAGGCC-----AAAGTG 1358
QY 244 LysLeuValAspPheAlaHis-----ValAlaGluGly----- 254
Db 1359 TGGATGATGACTTTGGGAAAAACCAAGCCCTGCTGAGGCCAGACCTTCGAGCATGAC 1418
QY 255 -----AspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIle 270
```

```
Db 1419 GTCCCTGGCAGGAGGGGAACCGGAGGATGGCTACCTCTCGGGGCTCAATAACCTCGTC 1478
Qy 271 LysPheValSerAspIleValProGluThrPro-----HisThrGlnPro 285
Db 1479 GACATCCTGACCGAGATGTCACAGGATGCCCACTGCCTGAGCTGCCACCGCCCTCCCT 1538
Qy 286 LeuGlyPro 288
Db 1539 ---GGCCCC 1544

RESULT 10
US-09-902-540-1096/c
; Sequence 1096, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1096
; LENGTH: 14555
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(14555)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1096

Alignment Scores:
Pred. No.: 0.658 Length: 14555
Score: 108.00 Matches: 81
Percent Similarity: 34.29% Conservative: 15
Best Local Similarity: 28.93% Mismatches: 104
Query Match: 7.10% Indels: 81
DB: 4 Gaps: 12

US-10-042-894A-8 (1-289) x US-09-902-540-1096 (1-14555)
Qy 5 HisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeu 24
Db 2867 CATCGACCTGAAGCGGCTGTTCTCCGCCACGCGGCGACGGAGCCCGTCATGCGCGTTG 2808
Qy 25 Ile-AspGlySerGlyLeu----- 30
Db 2807 ATCCGATGCTGAGGGGCTGTCGCGGGGTTCCGCGTCCAGGCCACGAGGAGTGGTCCAGA 2748
Qy 31 ----PheTyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrG1 49
Db 2747 AGGTACCATGACCTGCTGGAGCTGGAGCGCGGAGGCTGGAGACGGAGCCCTCACCA 2688
Qy 49 uAlaPhe-----SerAlaHisAlaAlaValProAlaArgAlaArgAs 63
Db 2687 AGCTCTACCTGCGGCTGGGCGCGCACTGCACACCTGAAGGCGACGCGCTCGCTGG 2628
Qy 63 pThrPhePheProArgPheHis-----GlyThrArgLeuLeuProThrGluAl 79
Db 2627 GCATGCAGACCTGAGCGACATCGCCCAACAGCTGGAGGACGCGCTCGCCCCGCTGAAGG 2568
Qy 79 aGlnProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGly----- 96
Db 2567 CGAACCTCAGAAAGATG--CCGCGCGCGGTGGTGNACATCCTGCTTCAGCGGCTGGACCT 2510
Qy 97 -----PheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProPr 114
Db 2509 CTTCTGCTGCGCGGCGAGGCCACGCGGAGGGGCTGGGGAGGGCTTCCTCGGACCCCGGC 2450
```

```
Qy 114 oSerSerProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerVa 134
Db 2449 GTCCCGCTGGCGCAGCTGGTGGCGAGACCCCGCGCGCCCGCCAGAGGC----- 2401
Qy 134 lleuLeuGlyPheArgValSerGlyValArgValValGlyPro-----G1 149
Db 2400 -----CGGCGCCATGGTCCGCGCCACCGGGGGCTTGGCGCGCGCTCGCCGCGC 2354
Qy 149 uGlyAlaValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValAr 169
Db 2353 CGSTGCCGTC-----CGCGAGCCCGTCTCTGTCGCGACGCGCTGCGCGGAGTCCG 2303
Qy 169 g---ArgValLeuArgArgTyrValSerSerValAlaAlaAspGluGlyMetAspCys----- 186
Db 2302 CGTACGCGGCTGGCGCGTGTCTGTC-----GTGCGAGTGACGGCGCTGATCGCGGAGG 2249
Qy 187 -----AlaLeuAlaAlaAlaValTyrGlyLysGlyValLeuSerGlnLe 203
Db 2248 TGGAGCGCTGCGGAGGTGGCTGCGCGTGGAGGAGCGGGCGCGAACTGGA----- 2194
Qy 203 uArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSerAlaSerIl 223
Db 2194 ----- 2194
Qy 223 eLeuLeuGlyTyrAspAlaAla-----AlaValAlaAlaGlyGlyAspGlyGlyGly 240
Db 2193 -----CGGGTGCGCGTCTGCTGGCGAAGCAGGGGCTGCTGGCGGAGCGCGGAGGC 2140
```

## RESULT 11

```
US-09-248-796A-4664
; Sequence 4664, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4664
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4664
```

## Alignment Scores:

```
Pred. No.: 0.0114 Length: 954
Score: 107.50 Matches: 46
Percent Similarity: 37.08% Conservative: 20
Best Local Similarity: 25.84% Mismatches: 53
Query Match: 7.07% Indels: 59
DB: 4 Gaps: 9
```

US-10-042-894A-8 (1-289) x US-09-248-796A-4664 (1-954)

```
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 28 TTAAATTCACCTACCTCTTCTTAAACATCAAGCTGTGCGCAT----- 69
Qy 21 LeuGlyProLeuIleAspGly-----SerGlyLeuPheTyrLysProLeuGln 36
Db 70 -----GATGGCTGCCTTCAACCGATTTCATTATTATTATAATTGACTGTT 114
Qy 37 AlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAla 56
Db 115 -----CAGCAAGAAATTTGATTTCCTACTCACTCAAGTTTCAACTGCAC----- 153
```

```
QY 57 ValProAlaArgIleArgAspThr-----PhePheProArg 68
Db 154 ---GATCAAGTGTCAAGATGCTCCATTAGGTCCCAATTGTCCTGATGCCTACA 210
QY 69 PheHisGlyThr-----Arg 73
Db 211 TTCTATGGGACATTGACACAAGGAGATGTATCAAAACTCAAAACAGCAGGTCTGGATCA 270
QY 74 LeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAspLeu 93
Db 271 CTGTATCAGCAGAGACAGACAGATAA-----CAATACATTGTACTACTGAATTCG 324
QY 94 LeuAlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAla----- 109
Db 325 TACCATTGGCTTTACCATCTCTAGTATATTGACATCAAGTTGGCGCCAAATTGACTGAC 384
QY 110 -----IleThrTrpProSerProGluProTyrlleAlaLysCysLeuAlaLys 127
Db 385 GATGAAGTTACTGCTCCAGAGAAA-----ATTGCCAGTTTACAAAAGTC 429
QY 128 AspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145
Db 430 AGTGATTCAACTACTAGTGGAGTTTAAATTTTCAAGAAATTTGTGTATGAAAGTC 483

RESULT 12
US-09-252-991A-12872/c
; Sequence 12872, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12872
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12872

Alignment Scores:
Pred. No.: 0.012 Length: 984
Score: 107.50 Matches: 91
Percent Similarity: 35.86% Conservative: 32
Best Local Similarity: 26.53% Mismatches: 117
Query Match: 7.07% Indels: 105
DB: 4 Gaps: 15

US-10-042-894A-8 (1-289) x US-09-252-991A-12872 (1-984)
QY 8 GluHisGlnValAlaGlyHisArgAlaSer-----AlaSerLysLeuGlyPro 23
Db 968 GAAACCAATGGCGGCGGACGCTCTACGTGACCGGCCCATCTGGGAATACTGGCGGCGG 909
QY 24 LeuIleAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArgGlyGlu--- 42
Db 908 TCCGCGGACGCGGACGACAGTTT-----ACGCTACGCGTGTGATTTCCGCGTGTGT 855
QY 43 -----HisGluValAlaPheTyrlsGluAlaPhe----- 51
Db 854 CCACGTTGAAGCAGATTGAGGAGGACGCGCGCTCTCCGCGCTTCCGAGGCTGA 795
QY 52 -----SerAlaHisAlaAlaValPro-AlaArgIleAr 62
Db 794 TCCGCGCGTCCGCGACGCGCTCGTCTCGACGCGGAGCAATTCACGCTGTCTCCG 735
QY 62 AspThr-----PhePheProArgPheHisGlyThrArgLeuLeuProThrG1 78
```

```
Db 734 GCGTACGGCGGCGCAGTACCACTTCGCCGACGCTGGTGA-----ATGACTGGCG 684
QY 78 ualaGlnProGlyGluProHisProHisLeuValLeu----- 90
Db 683 AGAACAGCCCGGTGCGCGGCGCGGACCTCTGCTCGTCCAGCGCCGCGATACAGGC 624
QY 91 -----AspAspLeuLeuAlaGlyPheGlnAl 99
Db 623 GCGGCTTGGAGGCGCGAGATCTGATCGCGCGGACGATTTCTTGGCCGTATAGCAGC 564
QY 99 aProCysValAlaAsp----- 105
Db 563 CCTTCTTGAAGCTGACCCCGCGCGCTGCAGGTTGATCATCTGTGGGATGAACAGCT 504
QY 105 e-LysIleGlyAlaIleThrTrpProSerProSerProGluProTyrlleAlaLysCysL 125
Db 503 CGCGGTAGCGCGCAACACCTGGCG-----ATGCTGC-GGACACTGGCGGAGCAGC 451
QY 125 euAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgV 145
Db 450 CAATCTCGAGCGCGCTCGCGCTGTCTCGC-GGAGCGTGGCAGCACCGCTCGGC 392
QY 145 alValGlyProGluGlyAlaValTrpArgThrGluArgPro-----GluValLysAla-Met 163
Db 391 GCGTTGCGCGGAAACCATAGCTCGACGCGGCGTCCGCCAGGCGGCGGAGCAGGTC 332
QY 164 AspThrAlaGlyValArgValLeuArgArgTy-Val----- 176
Db 331 GCGTGGCGGCGCATCTCGCGCTCTCGCGCGGGTGTCTATCCCAAGGCGCGCAGTGC 272
QY 177 -----SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaAala 191
Db 271 CTGCTGGCGTCCGCGCAAGCTATGCGAGCCAGGC-----TGGCTCTCTCGTGGCAAG 218
QY 192 ValTyrlsGlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
Db 217 GCTGGCTTGGAGAAATACGCGGTATTTCTCAGGTCCGCAACTGGGCGCTC----- 167
QY 212 GluGlnThrLeuPheHisPheTyrlsAlaSerIleLeuLeuGlyTyrlsAlaAlaAala 231
Db 166 GAGCAGCTCGCGCGCATCCGACGAGCGGCGCTCGCGCTCGGCGGAGGATGCGGAAGCT 107
QY 232 -----ValAlaAlaGlyAspGlyGlyValThrValLys 244
Db 106 CGATAGCAGCGCGCTTGTATGTTGAGCGCGCGCGAGGCTGGAGGTCTCGTGTGAG 47
QY 245 LeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAsePheHisPheLeuGly 264
Db 46 -----GTAGGCGAGGTTGACAGGTTCAGCTGGCGCTGGAGAACTTGGC 5
QY 265 Gly 265
Db 4 GGC 2

RESULT 13
US-09-016-434-1132/c
; Sequence 1132, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
```



```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-348-006B-6

Alignment Scores:
Pred. No.: 0.251 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 1 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-348-006B-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
DB 3077 AGTTCTTGGCGAGACTTGTCCCGCAGGACGTCCTCGGTAGCGGA----- 3033

QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
DB 3032 ---CGGGGGGCTGAAGGGCCAGGGCCCGCGCGTGTGGCTCGCACTTGGAGGTGTCAT 2976

QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
DB 2975 AGCCCGTGTCCGGCTTCAGCGCCCTGACGCTGACGTGAGCGTCT----- 2936

QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
DB 2935 -----CCGCGCCCGCTCAGCCGGCTCGCGGAGCTCAGTCTCTCGGG 2892

QY 82 -----GlyGluProHisProHisLeuValLeuAspLeuLeuAlaG1 96
DB 2891 CAGGGCCCGAGGACCGCGCTCCGCGACGGCCCGCTGTATTGACGAT----- 2843

QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
DB 2842 -----GGCCCCGTGTGGCTCGCGCGCGCACCGGTGGCAGCAGCGGA 2802

QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
DB 2801 GAAGGAC-----GGTCCCGCGCGAGGCGTTCGCCGCGCGCTCCA 2763

QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
DB 2762 GAATCTCGGGTGGCCACGGGGCGTGTCTCCGGGATGCTCAGGACCTCGGCTGCCTCT 2703

QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluValLy 161
DB 2702 CGCCCGAGCGCCCGGCTCGCGGCGCGCAAGCCGGAACACATAGTGTGGCCCTTGTGCA 2643

QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
DB 2642 CGCTGTATGCGGTGACCGGTCTCCGAGAGCGGGAACTCCAGGTGGCCAGGGCGGTGCG 2583

QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
DB 2582 AGTCCTCAGCGCCAACTGACGAGCGGTAGCCAGCACCTGTCTCGCGGTGCCAGCG 2523

QY 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
DB 2522 GGGGCTCCACAGGTGCCAGCAGGTGCTCGGGGTCTGCTGCACCGA----- 2474

QY 206 uLysAlaTrpPheGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
DB 2473 -CAGGGTTGGGGCCCGACACTGCTCCCTTGTGTCAACAACAC-----CT 2430

QY 225 uGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
DB 2429 TGGGTTGTCTCGAGCGCCATCGCCCTTCATGTGTGTAGCGGCTACCGTGTAGGATGACG 2370

; SEQUENCE CHARACTERISTICS:
; LENGTH: 6000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-800-825A-6/c

Alignment Scores:
Pred. No.: 0.251 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 2 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-800-825A-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
DB 3077 AGTTCTTGGCGAGACTTGTCCCGCAGGACGTCCTCGGTAGCGGA----- 3033

QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
DB 3032 ---CGGGGGGCTGAAGGGCCAGGGCCCGCGCGTGTGGCTCGCACTTGGAGGTGTCAT 2976

QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
DB 2975 AGCCCGTGTCCGGCTTCAGCGCCCTGACGCTGACGTGAGCGTCT----- 2936

QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
DB 2935 -----CCGCGCCCGCTCAGCCGGCTCGCGGAGCTCAGTCTCTCGGG 2892

QY 82 -----GlyGluProHisProHisLeuValLeuAspLeuLeuAlaG1 96
DB 2891 CAGGGCCCGAGGACCGCGCTCCGCGACGGCCCGCTGTATTGACGAT----- 2843

QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
DB 2842 -----GGCCCCGTGTGGCTCGCGCGCGCACCGGTGGCAGCAGCGGA 2802

QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
DB 2801 GAAGGAC-----GGTCCCGCGCGAGGCGTTCGCCGCGCGCTCCA 2763

QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
DB 2762 GAATCTCGGGTGGCCACGGGGCGTGTCTCCGGGATGCTCAGGACCTCGGCTGCCTCT 2703

QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluValLy 161
DB 2702 CGCCCGAGCGCCCGGCTCGCGGCGCGCAAGCCGGAACACATAGTGTGGCCCTTGTGCA 2643

QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
DB 2642 CGCTGTATGCGGTGACCGGTCTCCGAGAGCGGGAACTCCAGGTGGCCAGGGCGGTGCG 2583

QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
DB 2582 AGTCCTCAGCGCCAACTGACGAGCGGTAGCCAGCACCTGTCTCGCGGTGCCAGCG 2523

QY 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
DB 2522 GGGGCTCCACAGGTGCCAGCAGGTGCTCGGGGTCTGCTGCACCGA----- 2474

QY 206 uLysAlaTrpPheGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
DB 2473 -CAGGGTTGGGGCCCGACACTGCTCCCTTGTGTCAACAACAC-----CT 2430

QY 225 uGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
DB 2429 TGGGTTGTCTCGAGCGCCATCGCCCTTCATGTGTGTAGCGGCTACCGTGTAGGATGACG 2370
```

```
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGCCCGTGTGGGCTTCAGCCCTTCAGCGTGAGCGTTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CCGGCCCGGCTCAGCGGGCTGCGGCAGCTCAGTCTCTCGGG 2892
QY 82 -----GlyGluProHisProHisLeuValLeuAspLeuAlaGl 96
Db 2891 CAGGGCCAGGGCAGCGCTCCGCGCCAGCCACCGTGATTGACGAT----- 2843
QY 96 yPheGlnAlaProCybValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
Db 2842 -----GGCCCCCTTGGCTTCGGCGGCACGGGTGGCAGCCAGCGGA 2802
QY 116 rProGluProTyrlleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
Db 2801 GAAGGAC-----GGTCCCGCCGAGGCGTTGCCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgVal----- 146
Db 2762 GAATCTGCGGGTGGCCACGCGGCTGCTCCTCGGATGCTCAGGACCTCGGCTGCCTCT 2703
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161
Db 2702 CGCCCCAGCGCGCGGCTCCGGGCGCGCAAGCGCGGAACATACGTGGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCCTGATGCCGTGTAGCGGTCTCTCGAGGGCGGGAACTCCAGGTGGCCAGGGGCGTGG 2583
QY 170 gValLeuArgArgTyTyValSerSerValAlaAsp-----GluGlyMetAspCy 186
Db 2582 AGTCCTCAGCGGCAAACTCGAGCGGTAGCCAGCACCTCGTCCTCCGCGGTGCCAGCG 2523
QY 186 sAlaLeuAlaAlaValTyrglyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGCTCCAGCGGTGCCAGAGGCTCCCTCGGGGGTCTGTGCACCGA----- 2474
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrsrAlaSerIleLeuLe 225
Db 2473 -CAGGGTTGGCGGCCAGCACCTGCTCCTTGTGTACAAACCAC-----CT 2430
QY 225 uGlyTyAspAlaAlaAlaValalaglyGlyAspGlyGly----- 239
Db 2429 TGGGTTTGTGAGGCGCATCGCCCTTCATGTGTAGGCGGCTACCGTGTAGGATAG 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValIl 258
Db 2369 CGGTCTCAGGCTGCAAGTTTGTGATGACCATCTCTGGGCATCGGCAGCATGACGTCT 2310
QY 258 eAspHisAsnPheLeuGlyLeuCyssSerLeuIleLysPheValSerAspIleValPr 278
Db 2309 TGATGCGCGGCGCGCGGCTCCGGCCCTCCATGCGCAGTAGTGCACCTGGTAGC 2250
QY 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CGCGGATCTGCGCGTGTGCGCGCGG 2222
```

Search completed: June 17, 2005, 06:38:38  
Job time : 230 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 01:52:53 ; Search time 615 Seconds  
(without alignments)

2781.797 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHTQPLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Qs/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111949\_8281/app\_query.fasta\_1.455  
-DB=N Geneseq 16Dec04 -QFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blastum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894 @CGN\_1\_470 @runat\_15062005\_111949\_8281 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_16Dec04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1521	100.0	1344	6 AAD43514	Aad43514 Maize ino
2	1485	97.6	922	6 AAD43513	Aad43513 Maize ino
3	1435	94.3	923	6 AAD43512	Aad43512 Maize ino
4	1406.5	92.5	1169	6 AAD43511	Aad43511 Maize ino
5	1307	85.9	3416	6 AAD43522	Aad43522 Maize ino

6	1233	81.1	899	6	AAD43518	Aad43518 Maize ino
7	808.5	53.2	643	6	AAD43519	Aad43519 Maize ino
8	742.5	48.8	1020	6	AAD43517	Aad43517 P. argent
9	722.5	47.5	1195	6	AAD43516	Aad43516 Eucalyptu
10	714.5	47.0	1130	3	AAC48750	Aac48750 Arabidops
11	700.5	46.1	1130	3	AAC40313	Aac40313 Arabidops
12	674.5	44.3	1104	3	AAC33685	Aac33685 Arabidops
13	674.5	44.3	1168	3	AAC39023	Aac39023 Arabidops
14	674.5	44.3	1243	3	AAC38692	Aac38692 Arabidops
15	644.5	42.4	1105	6	AAD43515	Aad43515 Soybean i
16	484	30.5	519	6	AAD43520	Aad43520 Maize ino
17	329	21.6	353	6	AAD43521	Aad43521 Maize ino
18	306.5	20.2	484	10	ADE82058	Ade82058 Arabidops
19	288	18.9	2930	4	ABL23306	Ab123306 Drosophil
20	273.5	18.0	464	6	ABL93575	Ab193575 Arabidops
21	265	17.4	876	4	ABL23307	Ab123307 Drosophil
22	201.5	13.2	1251	12	ADK65918	Adk65918 Human IPM
23	201.5	13.2	2212	8	AAD53176	Aad53176 Human kin
24	193	12.7	1341	11	ADI31483	Adi31483 Human cDN
25	192	12.6	2391	10	ADC06742	Adc06742 Human ino
26	192	12.6	2890	6	AAD38858	Aad38858 Human kin
27	191	12.6	2608	10	ADA53453	Ada53453 Human cod
28	189	12.4	1340	4	AAI60607	Aai60607 Human pol
29	189	12.4	1677	13	ADQ87184	Adq87184 Human tum
30	189	12.4	1737	5	AAH64793	Aah64793 Human sec
31	189	12.4	1737	10	ADC06746	Adc06746 Human ino
32	189	12.4	1739	4	AAH16813	Aah16813 Human cDN
33	189	12.4	1748	4	AAI58821	Aai58821 Human pol
34	189	12.4	1748	5	ADQ99042	Adq99042 DNA encod
35	189	12.4	1748	9	ADB48802	Adb48802 Novel hum
36	189	12.4	1757	5	AAH64800	Aah64800 Human sec
37	189	12.4	1791	11	ACN89795	Acn89795 Breat ca
38	189	12.4	1990	5	AAH64898	Aah64898 Human sec
39	183	12.0	1918	3	AACT7195	Aac77195 Human ORF
40	182.5	12.0	1222	10	ADC06748	Adc06748 Murine in
41	182.5	12.0	2306	10	ADF44508	Adf44508 Mouse kin
42	182	12.0	1863	8	ABX70876	Abx70876 Novel hum
43	178.5	11.7	1893	8	ABX70875	Abx70875 Novel hum
44	168	11.0	1723	10	ADC06744	Adc06744 Human ino
45	168	11.0	4461	6	ABR83730	Abk83730 Human cDN

#### ALIGNMENTS

RESULT 1  
AAD43514  
ID AAD43514 standard; DNA; 1344 BP.  
XX  
AC AAD43514;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 52..921  
FT /tag= a  
FT /product= "Maize IPPK protein #4"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX WPI; 2002-636540/68.  
DR P-PSDB; AAE26196.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
XX Claim 1; Page 64-65; 86pp; English.  
XX  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are used to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
SQ Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1,06e-130 Length: 1344  
Score: 1521.00 Matches: 289  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43514 (1-1344)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
DB 52 ATGTCGAGCTCCACCGCGGAGCAACCAAGTCGCGGCACCGGCCCTCCGCCAGCAG 111  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 112 CTGGGCGCGCTCATCGACGGCTCGGCCCTCTTCTACAGCGGCTCCAGCGCGGACCGT 171  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 172 GGGGAGCAGAGGTCCGCTCTATGAGCGGTCTCCGCCACCGCCCTCCCGCGCGC 231  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 232 ATCCGAGACACCTTCTTCCCGCGTTCCAGCGCAGCGACTCTCCCCACCGAGCGCAG 291  
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
DB 292 CCGGGGAGCGCATCTCCTCAGCTCGTCTCGACGACTCTCGCGGGTTTCAGGCGGCC 351  
QY 101 CysValAlaAspIleIleIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
DB 352 TCGCTCGCAGACATCAAGATCGCGCCATCATCGTGGCCACCGAGTTTCGCGGAGCCCTAC 411  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 412 ATCGCCAAAGTCTCGCCCAAGNCCCGCGGACACAGAGGTTCTGCTCGGATTCGCGTC 471  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
DB 472 TCGCGGTCCGAGTCTCGCGCCCGGAGGCGCGCTGTGGCGAGCGCGCGGAGGTG 531  
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180  
DB 532 AAGGCCATGGACACCGCGCGTCCCGCGTGTCTCGCGGTGTCTCGCGGTGTCTCGCGGTGCC 591  
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeu 200

DB 592 GACGAGGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTG 651  
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220  
DB 652 TCACAGCTCGCGAGCTCAAGCGGTGGTTCGAGGAGCAGACTCTGTTCACCTTACTCG 711  
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyGly 240  
DB 712 GGGTCGATCTTCTCGGCTATGATCTGCTGCAATCGGAGCGGAGATGGGGTGG 771  
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260  
DB 772 GTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGGGTGATGGGTGATTGACCAC 831  
QY 261 AnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280  
DB 832 AACTTCTCTGGCGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACATTGTTCCGGAGACT 891  
QY 281 ProHisThrGlnProLeuGlyProSer 289  
DB 892 CCTCATACGACGCCCTTGGGTCTTCT 918

RESULT 2  
AAD43513  
ID AAD43513 standard; DNA; 922 BP.  
XX  
AC AAD43513;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #3.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 53..922  
FT /tag= a  
FT /product= "Maize IPPK protein #3"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX WPI; 2002-636540/68.  
DR P-PSDB; AAE26195.  
XX  
PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 62-63; 86pp; English.  
XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are used to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA



Db 173 GGGGAGCAGAGTGGCTTTCTATAGGGGCTTCTCGCCCAAGCCGCTCCGGCCGC 232  
Qy 61 lIeArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 233 ATCCGAGACACCTTCTTCCCGCGGTTCACGCGCAGGACTCTCTCCCAACCGAGCGCAG 292  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
Db 293 CCGGGGAGCCGATCCGACCTCGCTCCGACGACCTCTCGCGGATTTAGGCGGCC 352  
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 353 TGGGTTCGAGACATCAAGATCGCGCCATCATGTGGCCACCGAGTTCGCGGAGCCCTAC 412  
Qy 121 lIeAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 413 ATCGCCAAAGTGGCTCGCCATGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCGCGTC 472  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 473 TCGCGCTCCGAGTCTGCTCCCGAGGCGCGGTGTGGCGACGAGCGCCCGAGGTC 532  
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
Db 533 AAGGTATGGACACCGTCGCGGTCCGCGGTGCTCCGCGGTACGTGTCATCCGGTTCG 592  
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValne 200  
Db 593 CGACGAGGGATGGACTGCGCGCTCGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTT 652  
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
Db 653 GTCACTGCTGCGGAGCTCAAGCGGTGTTCGAGGAGCGCTCTGTTCACCTTCTACTC 712  
Qy 220 rLaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyGlyGlyGly 240  
Db 713 GCGCTCGATTCCTCTGGGTATGATCTGCTCGAGTCGCGAGCGGAGGTGGGGTGG 772  
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
Db 773 GGTAAACAGTGAAGCTGGTGACTTTTGCCCATGTGGCCGAGGTGATGGGGTGATGAACA 832  
Qy 260 sLenPheLeuGlyGlyLeuCysSerLeuLysPheValSerAspIleValProGluTh 280  
Db 833 CAACTTCCTGGCGGCTCTGCTAGCTGATCAAGTTCTGTCATTTCTGACATTTTCCAGAG 892  
Qy 280 rProHisThrGlnProLeuGlyProSer 289  
Db 893 TCCTCAGACGACGCTTTGGGTCTTCT 920

RESULT 4

ID AAD43511 standard; DNA; 1169 BP.  
XX  
AC AAD43511;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #1.  
DE  
DE  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 84..806  
FT /\*tag= a  
FT /product= "Maize IPPK protein #1"  
XX  
XX  
PN WO200259324-A2.  
XX

PD 01-AUG-2002.  
XX 09-JAN-2002; 2002WO-US003120.  
XX 12-JAN-2001; 2001US-0261465P.  
XX (PION-) PIONEER HI-BRED INT INC.  
PA Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
PI P-PSDB; AAE26193.  
DR WPI; 2002-636540/68.  
XX P-PSDB; AAE26193.  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX Claim 1; Page 58-59; 86pp; English.  
PS The invention relates to novel inositol polyphosphate kinase (IPPK)  
XX polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
SQ Sequence 1169 BP; 238 A; 362 C; 347 G; 222 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.29e-120 Length: 1169  
Score: 1406.50 Matches: 277  
Percent Similarity: 96.21% Conservative: 2  
Best Local Similarity: 95.52% Mismatches: 10  
Query Match: 92.47% Indels: 2  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43511 (1-1169)

Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 84 ATGCCCGACCTCCACCCCGCGAGCACCAGTCGCGGTCCACCGCCTCCGCCAGCAG 143  
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 144 CTGGGCCCGCTCATCGACGGCTCCGCGCTCTTCTTACAAGCCGCTCCAGCGCGGACCGT 203  
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60  
Db 204 GGGGAGACGAGGTGCGCTTCTATAGGCGGTTCCTCGCCACGCGCGCTCCGGCCCGC 263  
Qy 61 lIeArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 264 ATCCGAGACACCTTCTTCCCGCGTTCCAGCGACGCGACTCTCTCCCAACCGAGCGCAG 323  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 324 CCGGGGAGCGCATCCGACCTCGCTCTCGACGACCTCTCTCGCGGGTTTGAGGCGGCC 383  
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 384 TCGGTTCGAGACATCAAGATCGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 443  
Qy 121 lIeAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 444 ATCGCAAGTACTCTCGCAAGGACCGCGGACACAGAGGCTTCTGCTCGGATTCGCGTC 503  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 504 TTGC--GTCCGAGTCTGTCGGCCCCGAGGCGCGGTGTGGCGGACGAGCGCCGAGGTG 561

QY 161 LysAlaMetAspThrAlaGlyValArgAtqValLeuArgArgTyrValSerSer-ValAl 180  
 DB 562 AAGGCTATGGACACCGTCGCGCGTCCGCGCGTCTCCGCGCTACGTGTATCCGCTTGC 621  
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValle 200  
 DB 622 CGACGAGGGATGAGTATCGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTT 681  
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
 DB 682 GTCAACAGTCCGCGAGCTCAAGCATGTGTGGAGGAGCAGACTCTGTTCACATCTTACTC 741  
 QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
 DB 742 GCGTTCGATTCTTCTGGGCTATGATGCTGTCAGTGCACAGCGAGCGAGGTGGGGTGG 801  
 QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
 DB 802 GGTAAACAGTGAAGCTGGTGACCTTGCCCATGTGGCCGAGGGGTGATGGGTGATTGACCA 861  
 QY 260 sAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
 DB 862 CAACTTCCTGGGCGAGCTCTGCTAGCTGATCAAGTTCTGACATTCTCCAGAGAC 921  
 QY 280 rProHisThrGlnProLeuGlyProSer 289  
 DB 922 TCCTTAGACGACGACCTTTGGGTCTCTCT 949  
 RESULT 5  
 ID AAD43522 standard; DNA; 3416 BP.  
 XX AAD43522;  
 AC AAD43522;  
 XX AAD43522;  
 DT 14-NOV-2002 (first entry)  
 XX Maize inositol polyphosphate kinase (IPPK) DNA #9.  
 DE Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic; gene; ds.  
 XX Zea mays.  
 XX Key Location/Qualifiers  
 FH 72..407  
 FT CDS /\*tag= a  
 FT /product= "Maize IPPK protein #6"  
 XX WO200259324-A2.  
 XX 01-AUG-2002.  
 XX 09-JAN-2002; 2002WO-US003120.  
 XX 12-JAN-2001; 2001US-0261465P.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 DR WPI; 2002-636540/68.  
 DR P-PSDB; AAE26201.  
 XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 XX Claim 1; Page 75-77; 86pp; English.  
 PS The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC

CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are used to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK DNA  
 XX  
 SQ Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.9e-110 Length: 3416  
 Score: 1307.00 Matches: 260  
 Percent Similarity: 90.34% Conservative: 2  
 Best Local Similarity: 89.66% Mismatches: 9  
 Query Match: 85.93% Indels: 20  
 DB: Gaps: 1

US-10-042-894A-8 (1-289) x AAD43522 (1-3416)  
 QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
 DB 72 ATGCCCGACCTCCACCCGCGGAGCACCAGTCCCGGTCCCGGCTCCCGCAGCAG 131  
 QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
 DB 132 CTGGGCCCACTCATCGACGACTTGGCTCTTCTACAGCCGCTCCAGGCCGCGACCGT 191  
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
 DB 192 GGGGAGCAGAGTCGCTTCTATGAGGGGTCTCCGCCCCACGCCGCTCCCGGCCGC 251  
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
 DB 252 ATCCGAGACACCTTCTTCCCGGTTCACGGCAGCGACTCTCTCCACCGAGGCCGAG 311  
 QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
 DB 312 CCGGGGAGCGCATCCGACCTCTCTCGACGACTCTCTCGGGGTTTGAGGCGCC 371  
 QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
 DB 372 TGGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 406  
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
 DB 407 -----ACCACGAGCGTCTCTCGGATTCGCGTC 436  
 QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
 DB 437 TCGGGGTTCGAGTCTCGGCCCGGAGGGCGCGGTGTGGGGAGCGAGCGCCGAGGTG 496  
 QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
 DB 497 AAGGCTATGGACATTCGCGCGTCCCGCGCTCTCCGCGCTACGTGTATCCGCTTGC 556  
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValle 200  
 DB 557 CGACGAGGGATGAGTATCGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTT 616  
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
 DB 617 GTCAACAGTCCGCGAGCTCAAGCGGTGTTCGAGGGGCGAGACTCTGTTCACATCTTACTC 676  
 QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
 DB 677 GCGTTCGATTCTTCTGGGCTATGATGCTGTCAGTGCACAGCGAGCGAGGTGGGGTGG 736  
 QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
 DB 737 GGTAAACAGTGAAGCTGGTGACTTTGCCCATGTGGCGGAGGGGTGATGGGTGATTGACCA 796  
 QY 260 sAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280

```

Db      797 CAACCTTCCTGGGGGGCTCTGCTAGCTGATCAAGTTTGTCTGACATTTGTTCCAGAGAC 856
QY      280 rProHisThrGlnProLeuGlyProSer 289
Db      857 TCCTCAGACGCGCCTTTGGGTCTTCT 884

RESULT 6
AAD43518
ID      AAD43518 standard; DNA; 899 BP.
AC      AAD43518;
XX      14-NOV-2002 (first entry)
DT      Maize inositol polyphosphate kinase (IPPK) DNA #5.
DE      Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW      nutritional value; animal feed; transgenic; gene; ds.
XX      Zea mays.
FH      Key
FT      Location/Qualifiers
FT      89..424
FT      /*tag= a
FT      /product= "Maize IPPK protein #4"
XX      WO200259324-A2.
PN      01-AUG-2002.
XX      09-JAN-2002; 2002WO-US003120.
PD      12-JAN-2001; 2001US-0261465P.
XX      (PION-) PIONEER HI-BRED INT INC.
XX      Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX      WPI; 2002-636540/68.
DR      P-PSDB; AAE26200.
XX      New inositol polyphosphate kinase polynucleotides and polypeptides,
PT      useful in modulating phytic acid biosynthesis by decreasing phytate or
PT      increasing non-phytate phosphorous to improve the nutritional value of
PT      animal feed.
XX      Claim 1; Page 73-74; 86pp; English.
XX      The invention relates to novel inositol polyphosphate kinase (IPPK)
CC      polypeptides and polynucleotides. Sequences of the invention are useful
CC      in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC      increasing non-phytate phosphorous to improve the nutritional value of
CC      animal feed, or to reduce the environmental impact of animal waste.
CC      Polynucleotides of the invention are to produce transgenic plants with an
CC      altered phenotype. IPPK proteins are used to screen compounds that
CC      modulate their activity and raising anti-idiotypic antibodies. The
CC      present sequence is maize IPPK DNA
XX
SQ      Sequence 899 BP; 140 A; 312 C; 276 G; 171 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2,44e-104      Length:      899
Score:          1233.00      Matches:      248
Percent Similarity: 88.42%      Conservative: 4
Best Local Similarity: 87.02%      Mismatches: 14
Query Match:    81.07%      Indels:      20
DB:             Gaps:      1

US-10-042-894A-8 (1-289) x AAD43518 (1-899)

QY      1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20

```

```

Db      89 ATGCCCGACCTCCACCCCGCGGAGCACCAAGTCGCCGCTCACCCGCCCTCCGCCAGCAAG 148
QY      21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db      149 CTGGGCCCACTCATCGACGGCTCTGGCCTCTTTACAAGCCGCTCCAGGCCCGCACCGT 208
QY      41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaAsp 60
Db      209 GGGGAGCAGAGGTGCGCTTCTATAGGGGTCTCCGCCACGCCGCCGCTCCGCCGCCG 268
QY      61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db      269 ATCCGAGACACCTTCTCCCGCGTTCCAGCGCACGCGACTCTCTCCACCGAGCGCAG 328
QY      81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db      329 CCCGGGAGCGCATCCGTACCTCGCTCGACGACCTCTCTCGGGGTTTGAGGGGCC 388
QY      101 CysValAlaAspIleLysIleGlyAlaIleThrTTPProProSerSerProGluProTyr 120
Db      389 TCGTTCGCAGACATCAAGATCGGTCCATCAGTACCATTG-AGCGAT----- 435
QY      121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db      436 -----CTGCTCGGATTCACGTC 453
QY      141 SerGlyValArgValValGlyProGluGlyAlaValTTPArgThrGluArgProGluVal 160
Db      454 TCCGGCGTCCGAGTCTGCGGCCCGAGGGCGCGGTGTGGCGAGCGAGCGCTCGAGGTG 513
QY      161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
Db      514 AAGGCTATGGACATTTGTCGGGTCCCGCGCTGCTCCGGCGTGCATGTTCATCCGTTGC 573
QY      180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyVal 200
Db      574 CGGCGAGGGGATGGACTGCGGCTCGCGCGCGGTGTACGGAGGAGGAGGTGGAGTCTT 633
QY      200 uSerGlnLeuArgGluLeuLysAlaTTPheGluGlnThrLeuPheHisPheTyrSe 220
Db      634 GTCACAGCTGCGGAGCTCAAGCGGTGGTTCCAGGGGCGAGACTCTGTTCACCTTCTACTC 693
QY      220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db      694 GCGCTCGATTCTTCTGGGCTATGATGCTGCTGCAGTCCGACGAGCGGAGGTGGGGTGG 753
QY      240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db      754 GGTAACAGTGAAGCTGGTGGACCTTGGCCCATGTGGCCGAGGGGTGATGGGGTGAATGACCA 813
QY      260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db      814 CAACCTTCCTGGGCGGGCTCTGCTAGCTGATCAAGTTTGTCTGACATTTGTTCCAGAGAC 873
QY      280 rProHisThrGln 284
Db      874 TCCTTAGACGCGAG 886

RESULT 7
AAD43519
ID      AAD43519 standard; DNA; 643 BP.
XX      AAD43519;
AC      AAD43519;
XX      14-NOV-2002 (first entry)
DT      Maize inositol polyphosphate kinase (IPPK) DNA #6.
DE      Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW      nutritional value; animal feed; transgenic; ds.
XX      Zea mays.

```

PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
XX WPI; 2002-636540/68.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 74; 86pp; English.  
XX  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
SQ Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;  
  
Alignment Scores:  
Pred. No.: 2,41e-65 Length: 643  
Score: 808.50 Matches: 162  
Percent Similarity: 93.14% Conservative: 1  
Best Local Similarity: 92.57% Mismatches: 10  
Query Match: 53.16% Indels: 3  
DB: 6 Gaps: 0  
  
US-10-042-894A-8 (1-289) x AAD43519 (1-643)  
QY 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerLys 20  
DB 117 ATCCCGGACCTCCACCGCGCGAGCACCAAGTCGCGGCTCACCAGCGCTCCGCGAG 176  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 177 CTGGGCGCGCTCATCGACGCTCGGCGCTTTCTACNAGCGCTCCAGGCGCGGACCGT 236  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 237 GGGGAGCAGCAGGTGCGCTTCTATAGGCGCTTCTCGGCCACCGCGCGCGCGCGC 296  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 297 ATCCGAGACACTTCTTCCCGCGGTTCCACGGCAGCGACTCTCCCGCCAGCGGCGAG 356  
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
DB 357 CCCGGGAGCGGATCCGACCTCTGCTCTGACGACCTCTCGCGGGTTTGAGGCGCC 416  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
DB 417 TCGTTCGACACATCAAGATCGGCGCCATCAGGTGGCCACCGAGTTCGCGGAGCCCTAC 476  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 477 ATCGNCAAGTACCTTNGCCAAAGACCGCGGACCCAGCGAGCGTTCGTCTCGATTCCGCGTC 536  
QY 141 SerGlyValArgValValGlyProGluGlyValAlaValTrpArgThrGluArg-ProGluVa 160

DB 537 TTGC-GTCCGAGTCGTGCGCCCGAGGCGCGCGTGTGGGAGCGAGCGCCCGGGGT 594  
QY 160 llysalamMetAspThr-AlaGlyValArgValLeuArg 173  
DB 595 GAANGCTATGGACACCCGTCGGNGNCCGCGGNGTGTCTCGG 635  
  
RESULT 8  
AAD43517  
ID AAD43517 standard; DNA; 1020 BP.  
XX  
AC AAD43517;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.  
XX  
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;  
KW nutritional value; animal feed; gene; ds.  
XX  
OS Parthenium argentatum.  
XX  
FH Key Location/Qualifiers  
FT CDS 21..908  
FT /tag= a  
FT /product= "P. argentatum IPPK protein #1"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
XX WPI; 2002-636540/68.  
DR P-PSDB; AAE26199.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 71-72; 86pp; English.  
XX  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is P. argentatum IPPK DNA  
XX  
SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 5,26e-59 Length: 1020  
Score: 742.50 Matches: 149  
Percent Similarity: 67.83% Conservative: 45  
Best Local Similarity: 52.10% Mismatches: 79  
Query Match: 48.82% Indels: 13  
DB: 6 Gaps: 5  
  
US-10-042-894A-8 (1-289) x AAD43517 (1-1020)  
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLeuGlyPro 23  
DB 24 CTCAAGCGCCGACATCATCAGGTTGCTGGACATGAAGCTGGGCTCGGGAAGCTTGGCCCA 83

QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43  
DB 84 CTCATGTGATGATTCAGCGCGGTTTACAAACACATGCGAGGTGATTAACCGTGGTCAGAA 143  
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAsp 63  
DB 144 GAAGTAGCCCTTTATGAATCTTTCTTCTAACAATAATATTCCAGAACACATACGC--- 200  
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83  
DB 201 AAATTCCTTCTATATATTATGCGACCAAAATCATG-----AAGCATCCACTGGCTCT 254  
QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103  
DB 255 GACCATCTCCATCATGGTGTGCAAGATCTTACATCAGCTCATGTCACCATCTGTATG 314  
QY 104 AspIleLeuIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123  
DB 315 GACATCAAAATCGGTCCAGAACATGGCGCCAGAAAGCTTCCGAGGCGTACATTGCAAAA 374  
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143  
DB 375 TGCCTAAAAAAGGATAGGGAAGACACAGTATTCCATTGGGATTCAGGATCTCCGGCTG 434  
QY 144 ArgValValGlyProGluGlyAla---ValTyrArgThrGluArgProGluValLysAla 162  
DB 435 CAAGTCATATCGATGATGGTCAGCGGTTTATAGCCTCATAGAAATTACATGGGTAAA 494  
QY 163 MetAspThrAlaGlyValArgArgValLeuArgTyrTyrValSerSerValAlaAspGlu 182  
DB 495 ACCGGCCAGCTGATGTTAGACTACTTCTTAGGAAATTTGTTCTTCTAACCGCTGCA 554  
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194  
DB 555 GAGATGGAATGCGCACAGCGCTAGCGCCGATTTCTTTAGCATCTTTTCTTTATGGT 614  
QY 195 GlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThr 214  
DB 615 GGGCTAATGGATATAGCTCACTGATGGAAATGGAACATGGTTTGAAGATCAACA 674  
QY 215 LeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAla 234  
DB 675 ATTTACCACCTCTATGCTTGTCTTTTGTTCATCTTTGAAAGAGGTTGGTTAAAA 734  
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly 254  
DB 735 GGTGCT---CGGTCAACCGCAGAAGTCAAACTATTGATTTTGTCTCATGTTACATGGT 791  
QY 255 AspGlyValIleAspHisAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSer 274  
DB 792 AATGGTGTATTGATCACAATTTCTTGGTGGGCTCTGTCTTTGATAAAGTTTCATTTCT 851  
QY 275 AspIleValProGluThr 280  
DB 852 GACATACTTTCGGAGACA 869  
RESULT 9  
ID AAD43516  
XX AAD43516 standard; DNA; 1195 BP.  
AC AAD43516;  
XX AAD43516;  
DT 14-NOV-2002 (first entry)  
XX Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.  
DE DE Eucalyptus grandis inositol polyphosphate kinase; IPPK; phytic acid; transgenic;  
KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;  
KW nutritional value; animal feed; gene; ds.  
XX Eucalyptus grandis.  
XX  
FH Key Location/Qualifiers

CDS 116..1048  
/\*tag= a  
/product= "E. grandis IPPK protein"  
XX WO200259324-A2.  
PD 01-AUG-2002.  
PF 09-JAN-2002; 2002WO-US003120.  
PR 12-JAN-2001; 2001US-026146SP.  
PA (PION-) PIONEER HI-BRED INT INC.  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX WPI; 2002-636540/68.  
DR P-PSDB; AAE26198.  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX Claim 1; Page 68-70; 86pp; English.  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are used to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is Eucalyptus grandis IPPK DNA  
XX Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 4,498-57 Length: 1195  
Score: 722.50 Matches: 149  
Percent Similarity: 66.31% Conservative: 38  
Best Local Similarity: 52.84% Mismatches: 84  
Query Match: 47.50% Indels: 11  
DB: 6 Gaps: 6  
US-10-042-894A-8 (1-289) x AAD43516 (1-1195)  
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23  
DB 119 CTCAAGGTCCCGATCATCAAGTCGCGGTCAACCGGGGAGACGGGGAAAGCTGGGGCA 178  
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43  
DB 179 CTGGTGGATGATTCGGGCGCTTCTATAAGCCTCTCCAGAGCGATCATCGCGGACACG 238  
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAsp 63  
DB 239 GAAATGGCCTTTTACGAGTCAITTCCTCAATACCGAGATCCAGGTCCACATTCG--- 295  
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83  
DB 296 AAATTCCTTCTCGCTTTCACGGAACCTAAGACTATT-----GAGCGCTCTGATGGATCG 349  
QY 84 ---ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysVal 102  
DB 350 GGTCTCTCAACCTCACCTGGTCTCTGAGGATCTCGTCTCGGTCGACGCAACCATCTCTC 409  
QY 103 AlaAspIleLeuIleGlyValaIleThrTrpProProSerSerProGluProTyrIleAla 122  
DB 410 ATGGACATCAAGACTGGATCCAGAACATGGTATTCGAGGCGCTCTGAGGAGTACATCAA 469  
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142





```
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.

PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2,29e-56 Length: 1130
Score: 714.50 Matches: 150
Percent Similarity: 66.55% Conservative: 39
Best Local Similarity: 52.82% Mismatches: 82
Query Match: 46.98% Indels: 13
DB: 3 Gaps: 8

US-10-042-894A-8 (1-289) x AAC48750 (1-1130)

QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerlysLeuGlyPro 23
Db 94 CTCRAAGGTCCTGACACCAAGTTCTGTCGTCACATTTGATGATGGGAAGCTCGGTCCA 153
QY 24 LeulleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 154 CTCGTAGATGACCAAGGCGGTTCTTCAAGCCACTTCAGGGAGATTCTCGTGGCGAACAC 213
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaAlaArgileArgAsp 63
Db 214 GAGGCTAAGTTCATGAGTCTTTTCATCGAACATGAGGTTCCAGATCACAATCCAT--- 270
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 271 AGATACTTCCCGGTGTATCAGGCACCTCAGCTAGTT-----GAAGCATCTCATGATCT 324
QY 84 ProHis---ProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysVal 102
Db 325 GGCAAGCTTCCTCATCTTGTCTTGTATGATGTTGTTTTCAGGTCAGCAAAACCCGTCGTA 384
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db 385 ATGGATGTTAGATTGGATCTAGGCATGTTACCGGATGATATCAGAGAATACTTCAAG 444
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 445 AAATGATTAAAGAAAGATAGACAGACCACCGGTTTCGTTGGGGTTCAGGGTTTCAGGT 504
QY 143 ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys 161
Db 505 TTTAAGATTTTGTATCACCAGAATCAAGTTTTCGAGAGCTGAGAAGAGCTTGTCTT 564
QY 162 AlaMetAspThrAlaGlyValArgValLeuArgValTyrValSer-----SerVal 179
Db 565 GGGTATAATGACGATGGTGTAGATTGGCTCTGAGGAAGTTTGTGTATCGAACTCTCC 624
QY 180 AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyLysGly 197
Db 625 GCTGACTCTAACTTCACACCAACATGCTTTTTCATCATCAGAGGTTTATGGCGTTGTAAC 684
QY 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHis 217
Db 685 GGGATCTTAGCGCAGTTTGTGAGCTTAAAGATTGTTTCGAAACCCAAACCGTTTACCAT 744
```



PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-01511080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	26-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	28-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
Alignment Scores:		
Pred. No.: 4,49e-55 Length: 1130		
Score: 700.50 Matches: 147		
Percent Similarity: 66.20% Conservatives: 41		
Best Local Similarity: 51.76% Mismatches: 83		
Query Match: 46.06% Indels: 13		
DB: 3 Gaps: 8		
US-10-042-894A-8 (1-289) x AAC40313 (1-1130)		
QY	4	LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db	94	CTCAAGGTCCTGAAACCAAGTTGCTGTCACATTGCTAGTGGGAAGCTAGGTCCA 153
QY	24	LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db	154	CTCGTAGATGACCAAGCGCGTCTTCAAGCCACCTTCAGGGAGATTCTCGTGGCGAACAC 213
QY	44	GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db	214	GAGGCTAAGTTCATGAGTCTTCACATCGAACAGAGGTTCCAGATCACATCCAT--- 270
QY	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db	271	AGATACTTCCCGGTGTATCACCGGCACCTCAGCTAGTT-----GAAGCATCTGATGATCT 324
QY	84	ProHis---ProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db	325	GGCAAGCTTCTCTCTTCTTCTGATGTTGTTTCAGGGTACGCCAACCCGTCGGTA 384
QY	103	AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db	385	ATGGATGTTAGATTGGATCTAGGACATGGTACCCGGATGATCAGAGAACTACTTCAAG 444
QY	123	LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db	445	AAATGTATTAAAGAAAGATAGACAGMCCACCACCGTTTCGTTGGGGTTTCAGGGTTTCAGGT 504
QY	143	ValArgValValGly---ProGluGlyAlaValTyrArgThrGluArgProGluValLys 161
Db	505	TTTAAGATTTTGTATCACCAGAAATCAAGTTTTTGGAGAGCTGGAAGAAGCTTGTCTTT 564
QY	162	AlaMetAspThrAlaGlyValArgArgValLeuArgTyrValSer-----SerVal 179
Db	565	GGGTATAATGCAGATGGTGTAGATTGGCTCTGAGGAAGTTTGTGTCTCGAACTCTCCC 624
QY	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 197
Db	625	GCTGACTCTAACTTGACACCAAACTGTGCTTTTGCATCAGAGGTTTATGGCGTTTGTAAAC 684
QY	198	GlyValLeuSerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHis 217
Db	685	GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGGTTCGAAACCCAAACGCTTTACCAT 744
QY	218	PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly 236
Db	745	TTCATTTCTGCTCGATCTCTGATGATTACGAGATGAATCAATCTTGATGAAGAGGAGA 804
QY	237	AspGlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAlaGlu 253
Db	805	GATGATGCACCGGCACACCGGCACACAGTGAAGCTGGTGAATTTCCGCGCATGTTCTTGAT 864
QY	254	GlyAspGlyValIleAspHisGlnPheLeuGlyLeuCysSerLeuIleLysPheVal 273
Db	865	GGAAACGGTGTATCGACCAATAATTTCTTGGGTGAGCTCTGCTCTTTTATAAAGTTTCATC 924
QY	274	SerAspIleVal 277
Db	925	AAAGATATTCTT 936
RESULT 12		
AAC33685		
ID AAC33685 standard; DNA; 1104 BP.		

XX AAC33685; 99US-0139462P.  
AC 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
DT 17-OCT-2000 (first entry) 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
DE 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
KW 23-JUN-1999; 99US-0140354P.  
XX 24-JUN-1999; 99US-0140695P.  
KW 24-JUN-1999; 99US-0140695P.  
XX 28-JUN-1999; 99US-0140823P.  
OS 29-JUN-1999; 99US-0140991P.  
XX 30-JUN-1999; 99US-0141287P.  
OS 01-JUL-1999; 99US-0141842P.  
PN 01-JUL-1999; 99US-0142154P.  
XX 02-JUL-1999; 99US-0142055P.  
PD 06-JUL-1999; 99US-0142390P.  
XX 08-JUL-1999; 99US-0142803P.  
XX 09-JUL-1999; 99US-0142920P.  
XX 12-JUL-1999; 99US-0142977P.  
XX 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 3949.  
AC Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS EP1033405-A2.  
PN 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 25-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134941P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.



PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.





PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0134768P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148119P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149310P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151130P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157753P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

```
PR 26-OCT-1999; 99US-01611359P.
PR 26-OCT-1999; 99US-01611360P.
PR 26-OCT-1999; 99US-01611361P.
PR 28-OCT-1999; 99US-01611820P.
PR 28-OCT-1999; 99US-01611992P.
PR 28-OCT-1999; 99US-01611993P.
PR 28-OCT-1999; 99US-01611993P.
PR 28-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1.27e-52 Length: 1243
Score: 674.50 Matches: 141
Percent Similarity: 65.36% Conservative: 42
Best Local Similarity: 50.36% Mismatches: 86
Query Match: 44.35% Indels: 11
DB: 3 Gaps: 7

US-10-042-894A-8 (1-289) x AAC38692 (1-1243)
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
DB 269 CTCAAAGTCCTGAACATCAGGTTCGAGACACATTGCTAAAGACGGGAAGCTGCTCT 328
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 329 CTCGTAGATGACAAAGGTCGGTTCTTCAAGCCACTTCAGGGCGATTCTCGTGTGAATC 388
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
DB 389 GAGGTAAAGTTCACGAATCTTCTCTCAACACAGAGGTTCCAGAACACATCCAT--- 445
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
DB 446 AGATATTTCCCGTGTATCAGGCATCTCAGCAGTT-----GAAGTCTGTGAGGCA 499
QY 84 ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
DB 500 -----GCCATGATGGTGTGGAAATCTTCTTGCAGAACTACTCAAAACCATCAGTAATG 553
QY 104 AspIleValGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
DB 554 GATGTTAAGATGGTTCGAAACATGGTATCTCTGATGCATCTGGAAGCCGAGAGGCTTCTTCGCGGG 613
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
DB 614 TGTTTGAAGAAACACCGGTACCAACACCGTCTCATCGGTTTCAGGATCTCTGTTTC 673
QY 144 ArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLysAla 162
DB 674 GAAGTGTATGATCACAAAGAAATCGAGTTTCTGGAAGCCGAGAGGAGCTTCTTCGCGGG 733
QY 163 MetAspThrAlaGlyValArgArgValLeuAtqArgTyrValSer-----SerValAla 180
DB 734 CTCGATGTAGATGGAGGAGATTGATCTCGAGGAAGTTTGTATCATCTAACTCATCTTCG 793
QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGly 198
DB 794 GACACTGGCTCGAAACCTGACTCTGCTTTGCTCGAGTGTTCACGGCGTTCACCGGG 853
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPhe 218
DB 854 ATCTTAACGCGATGTCTGGAACCTCAAGACCTGGTTCGAGAACCAACCGCTCTACCATTTTC 913
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGly 238
DB 914 AACTCTGTGTCGATTTTAATGGTCTATGAGAAATGAATCCATCTTGAAGGGAATGATGAT 973
QY 239 GlyGly---ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
DB 974 GATGCTAGACCACCAAGTCAAGCTGGTGGATTGCTCATGTTCTTGATGGTAAATGGTGC 1033
QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
DB 1034 ATTGACCATAACTTCTTTGGTGGTCTTTTGTCTTTTCATAAACTTCATTCGTGAGATTCTT 1093
```

```
RESULT 15
AAD43515
ID AAD43515 standard; DNA; 1105 BP.
XX
AC AAD43515;
DT 14-NOV-2002 (first entry)
XX
DE Soybean inositol polyphosphate kinase (IPPK) DNA.
XX
KW Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
nutritional value; animal feed; transgenic; gene; ds.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
CDS 12..851
FT /*tag= a
FT /product= "Soybean IPPK protein"
XX
FN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-026146SP.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
XX
DR P-PSDB; AAE26197.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 66-67; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is soybean IPPK DNA
XX
SQ Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 6.43e-50 Length: 1105
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservative: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 6 Gaps: 8
```

```
US-10-042-894A-8 (1-289) x AAD43515 (1-1105)
QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAsp 26
DB 24 CCGGAGCACACAGGTTCGGGGCAGCAAGGCGGAATCTCTGGGCCCACTCTGTCGAC 83
QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42
DB 84 GATTTTGGAAAATTTTACAAGCCCTTCAGACCCCAAAAGACGACGACACCCGCGCTCC 143
```

Qy	43	HisGluValAlaPheTyT	GluAlaPheSer	AlaHisAlaValProAla	ArgIleArg	62
Db	144	ACCGAACTCTCTTTT	CACCTCTCGCGCC	---GCCGCCACGACTACT	CTCCATCCGCG	200
Qy	63	AspThrPhePheProArg	PheHisGlyThrArg	LeuLeuProThr	GluAlaGlnProGly	82
Db	201	---TCTTCTTCCCGCTT	CACGGACACCGCCTCT	TGGACGCCTCCGACGCCT	CCGACGCCTCCGCT	257
Qy	83	GluProHisProHisLeu	ValLeuAspLeuLeuAla	GlyPheGlnAlaProCys	Val	102
Db	258	---CCCCACCCCTCACC	TGCTCTGGAGGACCTCT	CTCGGCTACTCCAAACCCCT	CCGCT	314
Qy	103	AlaAspIleIleGlyAla	IleThrTrpProProSer	SerProGluProFyrIle	Ala	122
Db	315	ATGACGCTAAGATCGCT	TCAGAACCTGGCCACCT	TGGGAGACTCCGAGGACTAC	ATATCTGC	374
Qy	123	LysCysLeuAlaLysAsp	ArgGlyThrThrSer	ValLeuLeuGlyPheArg	ValSerGly	142
Db	375	AAGTGCCTGAAGAAGG	CACAGAGTCCCTAGCTT	GCCCTGGGTTCAGAAATCT	CGGGA	434
Qy	143	ValArgValValGlyPro	GluGlyAlaValTrpArg	ThrGluArgProGluVal	LysAla	162
Db	435	GTCAAG-----GACT	CTATCTCTCTCTGGGAA	CCCTCACAGAAATCTCT	CCAGTGT	485
Qy	163	MetAspThrAlaGlyVal	ArgArgValLeuArgArg	TyrValSerSer	-----	178
Db	486	CTATCCGCCCATGGTGT	GCACCTGTTCTCAACAAG	TTCGTTCTCTTAATAATAT	CAAC	545
Qy	179	ValAlaAspGluGlyMet	AspCysAlaLeuAlaAla	ValTyrGlyLysGlyGly	198	
Db	546	CATGATGATCATCATCC	CGATTGCGCTTTCGCAAC	GGAGGTCTAC-----	GGCGCC	596
Qy	199	ValLeuSerGlnLeuArg	GluLeuLysAlaTrpPhe	GluGluGlnThrLeuPhe	HisPhe	218
Db	597	GTITTGAGCGCTTCGAC	AGCTCAAGACTGTTTCG	AGGTTCAACGGTGTAT	CACITC	656
Qy	219	TyrSerAlaSerIleLeu	LeuGlyTyrAspAlaAla	ValAlaGlyGlyAspGly	238	
Db	657	TATTCTTGTCTTGTGT	GTGTACGAGAGGAT-----	CTAGGGAAGGGAAGCT	710	
Qy	239	GlyGlyValThrValLys	LeuValAspPheAlaHis	ValAlaGluGlyAspGly	ValIle	258
Db	711	ACCAACCCCTCTGGT	CANAATCGTTGACTTTC	ACACGCTGTCGACGAAG	CGGTGTCTT	770
Qy	259	AspHisAenPheLeuGly	LysCysSerLeuIleLys	PheValSerAspIle	Val	277
Db	771	GATCACAATCTCTGGT	GGCTTTGTTCTTCATC	ACGTTCTCAAGATAT	CTCTA	827

Search completed: June 17, 2005, 04:11:12  
Job time : 625 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 01:57:18 ; Search time 4911 Seconds  
(without alignments)

2851.468 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLPGFS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111950\_8289/app\_query.fasta\_1.455  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894 @CGN 1\_1\_3731 @runat\_15062005\_111950\_8289 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521	100.0	1344	6 AX513570	AX513570 Sequence
2	1485	97.6	923	6 AX513568	AX513568 Sequence
3	1435	94.3	923	6 AX513566	AX513566 Sequence
4	1406.5	92.5	1169	6 AX513564	AX513564 Sequence

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	1307	85.9	3416	6	AX513583	AX513583 Sequence
6	1233	81.1	899	6	AX513578	AX513578 Sequence
7	1148.5	75.5	149142	8	AP005749	AP005749 Oryza sat
8	1148.5	75.5	157419	8	AP004772	AP004772 Oryza sat
9	1144.5	75.2	1570	8	AK072296	AK072296 Oryza sat
10	808.5	53.2	643	6	AX513580	AX513580 Sequence
11	742.5	48.8	1020	6	AX513576	AX513576 Sequence
12	722.5	47.5	1195	6	AX513574	AX513574 Sequence
13	718.5	47.2	25054	8	LUS310150	AJ310150 Linum usi
14	714.5	47.0	903	8	AT243592	AJ243592 Arabidops
15	714.5	47.0	903	8	AY072621	AY072621 Arabidops
16	714.5	47.0	903	8	AY147936	AY147936 Arabidops
17	714.5	47.0	1166	8	AF412073	AF412073 Arabidops
18	714.5	47.0	1170	8	ATH245521	AJ245521 Arabidops
19	714.5	47.0	57246	8	AB010069	AB010069 Arabidops
20	704.5	46.3	1130	8	AY085862	AY085862 Arabidops
21	674.5	44.3	1243	8	AY087217	AY087217 Arabidops
22	674	44.3	84869	12	AY350714	AY350714 Olimarabi
23	673.5	44.3	905	8	BT000196	BT000196 Arabidops
24	673.5	44.3	1390	8	AY136378	AY136378 Arabidops
25	673.5	44.3	116763	8	ATT211	AL163912 Arabidops
26	666.5	43.8	858	8	BT011900	BT011900 Arabidops
27	664	43.7	115040	8	AC124954	AC124954 Medicago
28	658.5	43.3	861	8	AY147935	AY147935 Arabidops
29	658.5	43.3	1068	8	ATH404678	AJ404678 Arabidops
30	648	42.6	130078	2	AC127018	AC127018 Sequence
31	644.5	42.4	1105	6	AX513572	AX513572 Sequence
32	469.5	30.9	43428	8	AP004981	AP004981 Lotus cor
33	464	30.5	519	6	AX513581	AX513581 Sequence
34	329	21.6	353	6	AX513582	CO604442 Sequence
35	288	18.9	2930	6	CO604442	CO604442 Sequence
36	288	18.9	85095	3	AC004573	AC004573 Drosophil
37	288	18.9	142257	2	AC017491	AC017491 Drosophil
38	288	18.9	182726	3	AC008002	AC008002 Drosophil
39	288	18.9	307741	3	AE003589	AE003589 Drosophil
40	286	18.8	1219	3	AY119207	AY119207 Drosophil
41	265	17.4	876	6	CO604443	CO604443 Sequence
42	217.5	14.3	3003	3	AK116726	AK116726 Ciona int
43	213.5	14.0	2136	5	BC081124	BC081124 Xenopus l
44	201.5	13.2	1251	6	CQ771531	CQ771531 Sequence
45	201.5	13.2	1251	9	AF432853	AF432853 Homo sapi

## ALIGNMENTS

RESULT 1	AX513570	Sequence 7 from Patent WO02059324.	1344 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	AX513570	Sequence 7 from Patent WO02059324.				
DEFINITION	AX513570					
ACCESSION	AX513570					
VERSION	AX513570.1	GI:23559670				
KEYWORDS	Zea mays					
SOURCE	Zea mays					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD					
REFERENCE	1	Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.				
AUTHORS		Novel inositol polyphosphate kinase genes and uses thereof				
TITLE		Patent: WO 02059324-A 7 01-AUG-2002;				
JOURNAL		PIONEER HI-BRED INTERNATIONAL, INC. (US)				
FEATURES	source	Location/Qualifiers				
		1..1344				
		/organism="Zea mays"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:4577"				
		52..921				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAD52954.1"				
		/db_xref="GI:23559671"				
		/translation="MSDLHPPEHQVAGHRASAKGLIDGSLFYKPLQAGDRGEHE				

VAFYEAFSAHAAPARIDTFPRPHGTLLPTEAQPGEPHPLVLDLLAGFAQPCV  
ADIKGAIWTPSSPPPIAKCLAKDRGTTSVLLGFRVSGVRVWPEGAWRTPEV  
KAMDYAGRRVLRVYSSVADEGMDCALAAVYGGKGVLSQLRELKAWFEQTLPHF  
YSASILLGYDAAVAAGGGGGVTVKLVDFAHVAEGDGVLDHNFLLGGLSLIKFVSDI  
VPETPTQPLGPS"

ORIGIN

Alignment Scores:  
Pred. No.: 9,21e-106 Length: 1344  
Score: 1521.00 Matches: 289  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AX513570 (1-1344)

QY	1	MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys	20
DB	52	ATGTCGCACTCCACCGCGGAGCACCAAGTCGCCGGCCACCGCGCTCCCGCCAGCAAG	111
QY	21	LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg	40
DB	112	CTGGGCGCGCTCATCGACGGCTCGGCGCTTCTACAAAGCGCTCCAGGCGGCGACCGT	171
QY	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60
DB	172	GGGAGACAGAGGTGCGCTTCTATGAGCGTTCCTCGGCCACGCGCGCTCCGGCGCGC	231
QY	61	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80
DB	232	ATCCGAGACACCTTCTTCCCGCGTTTCCAGCGCACCGGACTCTCTCCCAACCGAGGCGCAG	291
QY	81	ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro	100
DB	292	CCGGGGAGCGCATCTCACCTCGCTCTCGACGACTCTCTCGCGGGTTTCAGGCGCC	351
QY	101	CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr	120
DB	352	TGGTTCGCGAGACATCAAGATCGCGCGCATCATCGTGGCCACCGAGTTTCGCGGAGCCCTAC	411
QY	121	IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
DB	412	ATCGCCAAAGTGCTCCGCCAAGNACCGCGGAGCACACGAGGTTCGTCTCGATTCGCGTC	471
QY	141	SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal	160
DB	472	TCCGGCGTCCGAGTCGTGCGCCCGGAGGCGCGTGTGCGGACGCGGCGCGGAGGTG	531
QY	161	LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla	180
DB	532	AAGGCCATGGACACCGCGCGGTTCGCCCGGTGCTCCGCGGTACGTGTATCCGTTGCC	591
QY	181	AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu	200
DB	592	GACGAGGGATGGACTGTGCGTTCGCGCGGGGTGTACGGAGGAAGAGTGGAGTCTTG	651
QY	201	SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer	220
DB	652	TCACAGCTGCGGAGCTCAAGCGGTGGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCG	711
QY	221	AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly	240
DB	712	GGGTGCAATTCCTTGGGGCTATGATGCTGTCGAGTCGACGAGCGGAGATGGGGTGGG	771
QY	241	ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis	260
DB	772	GTGACGGTGAAGCTGTGTGACTTTGGCCATGTGGCCGAGGGTGTATGGGGTATTGACAC	831
QY	261	AsnPheLeuGlyLysCysSerLeuIleLysPheValSerAspIleValProGluThr	280
DB	832	AACTTCTCGGGGGGCTGCTCGCTGATCAAGTTCGTTCTGACATTGTTCCGGAGACT	891

QY	281	ProHisThrGlnProLeuGlyProSer	289
DB	892	CCTCATACGACGCTTTGGGTCTCTTCT	918
RESULT 2			
AX513568			
LOCUS	AX513568	923 bp	DNA linear PAT 05-OCT-2002
DEFINITION	Sequence 5 from Patent WO02059324.		
ACCESSION	AX513568		
VERSION	AX513568.1	GI:23559668	
KEYWORDS			
SOURCE	Zea mays		
ORGANISM	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1		
AUTHORS	Shi, J., Beach, L.R., Wang, H., Rafalecki, J.A. and Cahoon, R.E.		
TITLE	Novel inositol polyphosphate kinase genes and uses thereof		
JOURNAL	Patent: WO 02059324-A 5 01-AUG-2002;		
	PIONEER HI-BRED INTERNATIONAL, INC. (US)		
FEATURES	Location/Qualifiers		
source	1..923		
	/organism="Zea mays"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4577"		
CDS	53..922		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD52953.1"		
	/db_xref="GI:23559669"		
	/translation="MPDLHPHEHQVAGHRASAKPGPLIDGSLFYKPLQADRGHEH VAFYDAFSAHAAPARIDTFPRPHGTLLPTEAQPGEPHPLVLDLLAGFAQPCV ADIKGAIWTPSSPPPIAKCLAKDRGTTSVLLGFRVSGVRVWPEGAWRTPEV KAMDYAGRRVLRVYSSVADEGMDCALAAVYGGKGVLSQLRELKAWFEQTLPHF YSASILLGYDAAVAAGGGGGVTVKLVDFAHVAEGDGVLDHNFLLGGLSLIKFVSDI VPETPTQPLGPS"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.1e-103	Length:	923
Score:	1485.00	Matches:	283
Percent Similarity:	98.27%	Conservative:	1
Best Local Similarity:	97.93%	Mismatches:	5
Query Match:	97.63%	Indels:	0
DB:	6	Gaps:	0
	US-10-042-894A-8 (1-289) x AX513568 (1-923)		

QY	1	MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys	20
DB	53	ATGCCCGGACCTCCACCGCGGAGCACCAAGTCGCCGGCTCACCGCGCTCCCGCCAGCAAG	112
QY	21	LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg	40
DB	113	CCGGGCGCGCTCATCGACGGCTCCGGCTCTTCTACAAAGCGCTTCAGGCGGCGGACCGT	172
QY	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60
DB	173	GGGAGACGAGGTGCTTCTATGAGCGGTTCCTCGGCCACGCGCGCTCCCGCGCGC	232
QY	61	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80
DB	233	ATCCGAGACACCTTCTCCCGGTTCCACGCGACCGGACTCTCTCCCAACCGAGGCGCAG	292
QY	81	ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro	100
DB	293	CCCGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCTCGGGATTTCAGGCGGCC	352
QY	101	CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr	120
DB	353	TGCGTCGAGACATCAAGATCGGCGCATCACGTGGCCACCGAGTTCGCGGAGCCCTAC	412

Qy	121	IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
Db	413	ATCGCAAGTGCCTCGCCATGAGCGCGGGACCAACGAGCGCTTCTGCTCGGATTCCGGGTC	472
Qy	141	SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal	160
Db	473	TCGGGGCTCGAGTCGTGCGCGCCGAGGGCGCGTGTGGCGGACGGAGCGCCGGAGGTG	532
Qy	161	LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla	180
Db	533	AAGCCCATGACACCGCGCGCGTCCGCGCGTGTCTCCGGCGCTACGTTGTCATCCGTTGCC	592
Qy	181	AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeu	200
Db	593	GACGAGGGATGACTGTGCGCTCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTG	652
Qy	201	SerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer	220
Db	653	TCACAGCTGCGCGAGCTCAAGGGCGTGGTTCGAGAGCAGACTCTGTTTCCACTTCTACTCG	712
Qy	221	AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyGly	240
Db	713	GGGTGCATTCTTCTGGGCTATGATGCTGCTGCAGTCGACAGCGGAGGTGGGGTGGG	772
Qy	241	ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis	260
Db	773	GTACGGTGAAGCTGTGGACTTTGCCCATGTGGCCGAGGGTGTATGGGTGATTGACCAC	832
Qy	261	AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr	280
Db	833	AACTTCTCGCGCGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACATTGTTCCAGAGACT	892
Qy	281	ProHisThrGlnProLeuGlyProSer	289
Db	893	CCTCAGACCGAGCCTTTGGGTCTCTCT	919
RESULT 3			
LOCUS	AX513566	923 bp	DNA linear PAT 05-OCT-2002
DEFINITION	Sequence 3 from Patent WO02059324.		
ACCESSION	AX513566		
VERSION	AX513566.1	GI:23559666	
KEYWORDS			
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	1		
AUTHORS	Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.		
TITLE	Novel inositol polyphosphate kinase genes and uses thereof		
JOURNAL	Patent: WO 02059324-A 3 01-AUG-2002;		
FEATURES	PIONEER HI-BRED INTERNATIONAL, INC. (US)		
source	Location/Qualifiers		
	1..923		
	/organism="Zea mays"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4577"		
CDS	53..736		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD52952.1"		
	/db_xref="GI:23559667"		
	/translation="MDPLHPPEHQVAGHRASAKPGPLIDGSLFYKPLQAGDRGEHE		
	VAFYFAFAHAAPARIRDTFFPRPHGLTLLPTEAQGEPHPLVLDDLLAGFAPCV		
	ADIKIGAITWPPSSPEYIAKCLAMDRTTSLIGFRVSGVRVVPGEAVWRTERREV		
	KAMDVGVRVLRVRYSSACRGDGLRARGGVRRKRWSLVTAARAQGVVRGAASVPL		
	LLGVDSGL"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.85e-99	Length:	923
Score:	1435.00	Matches:	277

Percent Similarity:		96.21%	Conservative:	2
Best Local Similarity:		95.52%	Mismatches:	10
Query Match:		94.31%	Indels:	1
DB:		6	Gaps:	0
US-10-042-894A-8 (1-289) x AX513566 (1-923)				
Qy	1	MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys	20	
Db	53	ATGCCCGACTTCACCCGCGGAGACCAAGTCGCGGTCAACCGCGCTCCGCCAGCAAG	112	
Qy	21	LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg	40	
Db	113	CCGGGCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGT	172	
Qy	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60	
Db	173	GGGGAGCAGAGGTCTCTTCTATGAGGGCTTCTCCGCCACCGCGCGCTCCGCCCGC	232	
Qy	61	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80	
Db	233	ATCCGAGACACTTCTTCCCGCGTTCACGGCACCGGACTCTCTCCCGCCAGGCGCAG	292	
Qy	81	ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro	100	
Db	293	CCCGGGAGCGCATCCGACCTCGTCTCGACAGCTCTCTCGGGATTTGAGGCGGCC	352	
Qy	101	CysValAlaAspIleLysIleGlyAlaIleThrTyrProSerProGluProTyr	120	
Db	353	TGCGTCGACAGACATCAAGATCGCGGCATCAGCTGCGCACCGAGTTCCCGGAGCCCTAC	412	
Qy	121	IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140	
Db	413	ATGCCAAGTCCCTCGCCATGGACCGCGGACCAACGAGCTTCTGCTCGGATTCGGGTC	472	
Qy	141	SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal	160	
Db	473	TCGGCGTCCGAGTCTGCTGTCGCCGAGGCGCGTGTGGCGGACGGAGCGCCCGAGGTG	532	
Qy	161	LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAl	180	
Db	533	AAGGCTATGGACACCTCGCGCTCCCGCGCTCTCCGGCGCTACGTTGCTCCGCTGC	592	
Qy	180	aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLe	200	
Db	593	CGACGAGGGATGAGACTGCGCGCTCGCGCGCGGCGGTACGAGAGAAAGGTGGAGTCT	652	
Qy	200	uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe	220	
Db	653	GTCACTGCTGCGCGAGCTCAAGGCGTGGTTTCGAGGAGCAGCTCTGTTCCACTTCTACTC	712	
Qy	220	rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyLysAspGlyGly	240	
Db	713	GGCGTCGATTCTTCTGGGCTATGATGCTGTCAGTCCGACGAGCGGAGGTGGGGTGG	772	
Qy	240	yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi	260	
Db	773	GGTAAACAGTGAAGCTGGTGACTTTGCCATGTGGCGAGGTGATGGGTGATTGACCA	832	
Qy	260	sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr	280	
Db	833	CAACTTCTCCGGCGGCTCTGCTAGCTGATCAAGTTCTGTTCTGACATTGTTCCAGAGAC	892	
Qy	280	rProHisThrGlnProLeuGlyProSer	289	
Db	893	TCCTCAGACGACGCTTTGGGTCTCTCT	920	
RESULT 4				
LOCUS	AX513564	1169 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO02059324.			
ACCESSION	AX513564			
VERSION	AX513564.1	GI:23559664		

```
KEYWORDS
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1
AUTHORS     Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE       Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL     PATENT: WO 02059324-A 1 01-AUG-2002;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES    Location/Qualifiers
             1..1169
             /organism="Zea mays"
             /mol_type="unassigned DNA"
             /db_xref="taxon:4577"
             84..806
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="CAD52951.1"
             /db_xref="GI:23559665"
             /translation="MPDLPPHPQVAGHRASAKLGPLIDGSLFYKPLQAGDRGEHE
             /VAFYAFSAHAAPVIRIDTFPPFHGTRLLPTEAQPGPHPLVLDLILAGFEAPCV
             ADIKGAIPTWSPSPPIYAKLAKDRTTSLVLLGPRVLRPSRRGRVADGAPGGE
             GYGHRRPRPALRVLIRLPTRGWTARSRRCTEEKVESCHSCASSRHGWRSLCSTS
             TRRRFWAMLLQSQAEVGVG"
ORIGIN
Alignment Scores:
Pred. No.:      3,46e-97      Length:      1169
Score:          1406.50      Matches:    277
Percent Similarity: 96.21%    Conservative: 2
Best Local Similarity: 95.52%  Mismatches: 10
Query Match:    92.47%      Indels:     2
DB:             6           Gaps:       0
US-10-042-894A-8 (1-289) x AX513564 (1-1169)
QY      1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB      84 ATGCCCGACTCCACCGCGGAGCACCAGTCCGCGGTCCAGCGCCCTCCGCGCAGCAAG 143
QY      21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLeuProLeuGlnAlaGlyAspArg 40
DB      144 CTGGGCGCGGTCTATCGACGGCTCCGCGCTCTTCTACAGCGGCTCCAGCGCGGCGACCGT 203
QY      41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB      204 GGGGAGCAGAGGTGCGCTTCTATGAGCGTTCTCGCCACGCGCCCGTCCCGGCGCGC 263
QY      61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB      264 ATCCGAGACACCTTCTTCCCGCGTTCCAGCGCAGCGACTCTCTCCACCGCGGCGCAG 323
QY      81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB      324 CCGGGGAGCGGATCCGACCTTCGTCTCGACGACTCTCTCGGGGTTTGGGCGGCC 383
QY      101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB      384 TCGCTCGCAGACATCAAGATCGCGCCATCAGTCGCGCACCGAGTTCGCGGAGCCCTAC 443
QY      121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB      444 ATGCCCAAGTACCTCGCCCAAGACCGCGGAGCCACGAGCGTTCTGCTCGGATTCGCGCTC 503
QY      141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB      504 TTGC--GTCCGAGTCGTGCGGCCCGGAGGCGCGTGTGGCGGACGAGCGCGCGAGGTG 561
QY      161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
DB      562 AAGGCTATGACACCGCTCGGCGTCCGCGGTCTCTCGGCGTCTACGTGTCTCCGCTTGC 621

KEYWORDS
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLe 200
DB      622 CGACGAGGGATGGACTCGCGCTCGCGCGCGGTGTACGAGGAGGAGGAGGAGTCTT 681
QY      200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
DB      682 GTACAGCTGCGCGAGCTCAGGCTGTTGGAGGAGGAGACTCTGTTCCATCTTACTTC 741
QY      220 rLaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyG 240
DB      742 GCGCTCGATTCTTCTGGGCTATGATGCTGCTCAGTCCGAGCGGAGGAGGAGGAGTGG 801
QY      240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValLeAspHi 260
DB      802 GGTAAACAGTGAAGCTGGTGGACTTTGCCCATGTGCCGAGGAGTGTGGGTGATTGACCA 861
QY      260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
DB      862 CAACCTTCCTGGGCGAGCTCTGCTAGTGNATCAAGTTCGTTCTGACATTTGTTCCAGAG 921
QY      280 rProHisThrGlnProLeuGlyProSer 289
DB      922 TCCTTAGACGCGAGCCTTTGGGTCTTCT 949
RESULT 5
AX513583
LOCUS      AX513583
DEFINITION Sequence 20 from Patent WO02059324.
ACCESSION AX513583
VERSION   AX513583.1 GI:23559684
KEYWORDS  Zea mays
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
AUTHORS   Novel inositol polyphosphate kinase genes and uses thereof
TITLE     Patent: WO 02059324-A 20 01-AUG-2002;
JOURNAL   PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
            1..3416
            /organism="Zea mays"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4577"
            72..407
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD52959.1"
            /db_xref="GI:23559685"
            /translation="MPDLPPHPQVAGHRASAKLGPLIDGSLFYKPLQAGDRGEHE
            /VAFYAFSAHAAPVIRIDTFPPFHGTRLLPTEAQPGPHPLVLDLILAGFEAPCV
            ADIKGAI"
ORIGIN
Alignment Scores:
Pred. No.:      3,95e-89      Length:      3416
Score:          1307.00      Matches:    260
Percent Similarity: 90.34%    Conservative: 2
Best Local Similarity: 89.66%  Mismatches: 9
Query Match:    85.93%      Indels:     20
DB:             6           Gaps:       1
US-10-042-894A-8 (1-289) x AX513583 (1-3416)
QY      1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB      72 ATGCCCGACTCCACCGCGGAGCACCAGTCCGCGGTCCGCGCTCCGCGCAGCAAG 131
QY      21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
```



```
Db 132 CTGGGCCCATCTATCGACGACTCTGGCCCTCTTCTACAAGCCGCTCCAGCGCGCGACCGT 191
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 192 GGGAGACAGAGTGGCTTCTATAGAGCGTTCTCCGCCACCGCCGCTCCCGGCCGC 251
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 252 ATCCGAGACACCTTCTCCCGCGTCCACGGCAGCGACTCTCCACCGAGGCGCAG 311
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 312 CCGGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCCGGGGTTTGAGGGGCC 371
Qy 101 CysValAlaAspIleGlyValIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 372 TGGGTGCGACACATCAAGATCGGTGCCATCACGTG----- 406
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 407 -----ACCACGAGCGTTCTGCTCGGATTCGCGTC 436
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 437 TCGGGCTCCGAGTCTGTCGCCCGCAGGGCGCGTGTGGCGGACGGAGCGCGGAGTG 496
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 497 AAGGCTATGGACATTTGTGCGGTCCCGCGCTGCTCCGGCGCTACGTTGTCGTTGC 556
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValle 200
Db 557 CGACGAGGGATGAGACTGCGCGCTCGCGCGCGGTGTACGAGGAAAGGTGAGTCTT 616
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 617 GTCACAGCTGCGCGAGCTCAAGCGTGTTCGAGGGGCGACACTCTGTTCCACTTCTACTC 676
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db 677 GGGCTCGATTCTTCTGGCTTATGATGCTGTCAGTCCAGCGAGCGGAGTGGGGTGG 736
Qy 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 737 GGTAAAGTAGTGGTGGACTTGTGCCATGTGGCGAGGTGTATGGGGTGTATGACCA 796
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 797 CAACTTCTGGCGGGCTCTGCTAGCTGATCAAGTTGTTTCTTGACATTGTTCCAGAGAC 856
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 857 TCCTCAGACGACGCTTTGGGTCTCTTCT 884

RESULT 6
AX513578 899 bp DNA linear PAT 05-OCT-2002
LOCUS
DEFINITION Sequence 15 from Patent WO02059324.
ACCESSION AX513578
VERSION AX513578.1 GI:23559679
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
Shi,J., Beach,L.R., Wang,H., Rafaleki,J.A. and Cahoon,R.E.
Novel inositol polyphosphate kinase genes and uses thereof
Patent: WO 02059324-A 15 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..899
```

## CDS

```
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
89..424
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52958.1"
/db_xref="GI:23559680"
/translation="MPDLHPPEHVOVAGHRASAKLGLIDGSLFVKPLQAGDRGEHE
VAFYEAFAASAHAAPARIRDTFFPRFHGTRLLPTEAQGPBPHPLVLDLDDLAGEAPCV
ADIKIGAIT"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 3 21e-84 Length: 899
Score: 1233.00 Matches: 248
Percent Similarity: 88.42% Conservative: 4
Best Local Similarity: 87.02% Mismatches: 14
Query Match: 81.07% Indels: 20
DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AX513578 (1-899)
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 89 ATGCCCGCCTCCACC CGCGGAGCACCAAGTCGCGGTCCCGCCTCCGCGCAGCAAG 148
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 149 CTGGGCCACTCATCGACGCTCTGGCCCTCTTCTACAAGCCGCTCCAGCGCGGACCGT 208
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 209 GGGAGACGAGTCCGCTTCTATAGGGGTTCTCGCCCGCAGCGCCGCTCCGCGCCGC 268
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 269 ATCCGAGACACCTTCTTCCCGCGTTCACGGCAGCGACTCTCTCCACCGAGGCGCAG 328
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 329 CCGGGGAGCGCATCCGTAACCTCGTCTCGACGACTCTCTCGCGGGTTTGAGGGCCCC 388
Qy 101 CysValAlaAspIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 389 TGGCTGCGACACATCAAGATCGGTGCCATCAGTGCATG-AGCGAT----- 435
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 436 -----CTGCTCGGATTCACGTC 453
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 454 TCGCGGTCCGAGTCTCGCGCCCGAGGGCGCGGTGTACGGAGGAGAAAGGTGAGTCTT 513
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 514 AAGCTATGACACATTGTCGCGCTCCCGCGGTCTCGCGCTGCTCATGTCTCGCTTGC 573
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValle 200
Db 574 CGCGAGGGGATGAGTCTCGCGCTCCCGCGCGGTGTACGGAGGAGAAAGGTGAGTCTT 633
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 634 GTCACAGCTCGCGAGCTCAAGCGGTGTTTCGAGGGGCGAGACTCTGTTCACCTTCTACTC 693
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyLysAspGlyGly 240
Db 694 GGGCTCGATTCTTCTGGCTTATGATGCTGTCAGTCCGACGAGCGGAGGTGGGGTGG 753
Qy 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 753 -----
```

Db	754	GGTAACAGTGAAGCTGGTGGACCTTGCCCATGTGGCCGAGGTGATGGGTGATTGACCA	813				/organism="Oryza sativa (japonica cultivar-group)"
Qy	260	saaanpheleuGlyGlyLeuCyserLeuilelyspheValserAapIleValproGluTh	280				/mol_type="genomic DNA"
Db	814	CAACTTCCTCGGGCGGCTCTGTAGCTGATCAAGTTTGTTCGTACATGTTCCAGAGAC	873				/cultivar="Nipponbare"
Qy	280	rProHisThrGln	284				/db_xref="taxon:39947"
Db	874	TCTTAGACGCAG	886				/chromosome="2"
Db	874	TCTTAGACGCAG	886				/clone="OSJNBa0047A17"
Db	874	TCTTAGACGCAG	886				join(2637..2780,3298..3360)
Db	874	TCTTAGACGCAG	886				join(2637..2780,3298..3360)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.1"
Db	874	TCTTAGACGCAG	886				/note="hypothetical ORF predicted by GENSCAN"
Db	874	TCTTAGACGCAG	886				this category is not included in IRGSP standard"
Db	874	TCTTAGACGCAG	886				complement(5864..9107)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				complement(join(5864..6755,8843..9107))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK102815"
Db	874	TCTTAGACGCAG	886				complement(join(6112..6755,8843..9017))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				/note="Rieske iron-sulfur protein contains EST(s): AU165272(C63256), AU068962(C51210) contains full-length cDNA(s): AK102815"
Db	874	TCTTAGACGCAG	886				/codon_start=1
Db	874	TCTTAGACGCAG	886				/product="putative ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial precursor"
Db	874	TCTTAGACGCAG	886				/protein_id="BAD26134.1"
Db	874	TCTTAGACGCAG	886				/db_xref="GI:49388912"
Db	874	TCTTAGACGCAG	886				/translation="MLRVRRRLTTTALWRPAAAGARGLAGSLPGDDEFRRPQR PRFVNDSPFFAASRGSTLVRNQDVLTELPAATVAVKPKPSAKIIVVEYNHRYQ PGDPSKRAFAYVLSSGGRFIVASLRLVLKFLVLSMSAKDLALASLEVLSSIEPG TTVTVKRGKVFIRRTTEEDINLANSVDIGSLRDPQDQDAERVKNPWLAVIGVCTHL GCILPNAAGDFGGWFCPCGSHYDISGRKRPAPYNLEVPYTSFLEENKLLIG"
Db	874	TCTTAGACGCAG	886				9904..10750
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				join(<9904..10067,10732..>10750)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				/note="start and end point are not identified"
Db	874	TCTTAGACGCAG	886				join(9904..10067,10732..10750)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				/note="predicted by GENSCAN etc."
Db	874	TCTTAGACGCAG	886				/codon_start=1
Db	874	TCTTAGACGCAG	886				/product="hypothetical protein"
Db	874	TCTTAGACGCAG	886				/protein_id="BAD26135.1"
Db	874	TCTTAGACGCAG	886				/db_xref="GI:49388913"
Db	874	TCTTAGACGCAG	886				/translation="MENATVTADLAPWSQQCRPVGEREEGGVAAAKKEGRQIQKE EKGGRRRSGSGIRKTM"
Db	874	TCTTAGACGCAG	886				complement(11599..12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				complement(<11599..>12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK109226"
Db	874	TCTTAGACGCAG	886				complement(11599..12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				/note="contains full-length cDNA(s): AK109226"
Db	874	TCTTAGACGCAG	886				non-coding transcript
Db	874	TCTTAGACGCAG	886				probably inactive due to including stop codon(s) in CDS"
Db	874	TCTTAGACGCAG	886				join(15870..16043,16479..16484)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.5"
Db	874	TCTTAGACGCAG	886				join(15870..16043,16479..16484)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.5"
Db	874	TCTTAGACGCAG	886				/note="hypothetical ORF predicted by GENESH"
Db	874	TCTTAGACGCAG	886				this category is not included in IRGSP standard"
Db	874	TCTTAGACGCAG	886				complement(18932..25995)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.6"
Db	874	TCTTAGACGCAG	886				complement(join(18932..19648,19740..19836,20081..20194,20280..20444,20690..20827,20923..20999,21093..21207,21277..21382,21703..21764,21859..21993,22149..22331,22880..22936,23369..23417,23504..23652,24724..24771,25583..25995))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.6"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK120851"

Db	754	GGTAACAGTGAAGCTGGTGGACCTTGCCCATGTGGCCGAGGTGATGGGTGATTGACCA	813				/organism="Oryza sativa (japonica cultivar-group)"
Qy	260	saaanpheleuGlyGlyLeuCyserLeuilelyspheValserAapIleValproGluTh	280				/mol_type="genomic DNA"
Db	814	CAACTTCCTCGGGCGGCTCTGTAGCTGATCAAGTTTGTTCGTACATGTTCCAGAGAC	873				/cultivar="Nipponbare"
Qy	280	rProHisThrGln	284				/db_xref="taxon:39947"
Db	874	TCTTAGACGCAG	886				/chromosome="2"
Db	874	TCTTAGACGCAG	886				/clone="OSJNBa0047A17"
Db	874	TCTTAGACGCAG	886				join(2637..2780,3298..3360)
Db	874	TCTTAGACGCAG	886				join(2637..2780,3298..3360)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.1"
Db	874	TCTTAGACGCAG	886				/note="hypothetical ORF predicted by GENSCAN"
Db	874	TCTTAGACGCAG	886				this category is not included in IRGSP standard"
Db	874	TCTTAGACGCAG	886				complement(5864..9107)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				complement(join(5864..6755,8843..9107))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK102815"
Db	874	TCTTAGACGCAG	886				complement(join(6112..6755,8843..9017))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				/note="Rieske iron-sulfur protein contains EST(s): AU165272(C63256), AU068962(C51210) contains full-length cDNA(s): AK102815"
Db	874	TCTTAGACGCAG	886				/codon_start=1
Db	874	TCTTAGACGCAG	886				/product="putative ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial precursor"
Db	874	TCTTAGACGCAG	886				/protein_id="BAD26134.1"
Db	874	TCTTAGACGCAG	886				/db_xref="GI:49388912"
Db	874	TCTTAGACGCAG	886				/translation="MLRVRRRLTTTALWRPAAAGARGLAGSLPGDDEFRRPQR PRFVNDSPFFAASRGSTLVRNQDVLTELPAATVAVKPKPSAKIIVVEYNHRYQ PGDPSKRAFAYVLSSGGRFIVASLRLVLKFLVLSMSAKDLALASLEVLSSIEPG TTVTVKRGKVFIRRTTEEDINLANSVDIGSLRDPQDQDAERVKNPWLAVIGVCTHL GCILPNAAGDFGGWFCPCGSHYDISGRKRPAPYNLEVPYTSFLEENKLLIG"
Db	874	TCTTAGACGCAG	886				9904..10750
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				join(<9904..10067,10732..>10750)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				/note="start and end point are not identified"
Db	874	TCTTAGACGCAG	886				join(9904..10067,10732..10750)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				/note="predicted by GENSCAN etc."
Db	874	TCTTAGACGCAG	886				/codon_start=1
Db	874	TCTTAGACGCAG	886				/product="hypothetical protein"
Db	874	TCTTAGACGCAG	886				/protein_id="BAD26135.1"
Db	874	TCTTAGACGCAG	886				/db_xref="GI:49388913"
Db	874	TCTTAGACGCAG	886				/translation="MENATVTADLAPWSQQCRPVGEREEGGVAAAKKEGRQIQKE EKGGRRRSGSGIRKTM"
Db	874	TCTTAGACGCAG	886				complement(11599..12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				complement(<11599..>12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK109226"
Db	874	TCTTAGACGCAG	886				complement(11599..12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				/note="contains full-length cDNA(s): AK109226"
Db	874	TCTTAGACGCAG	886				non-coding transcript
Db	874	TCTTAGACGCAG	886				probably inactive due to including stop codon(s) in CDS"
Db	874	TCTTAGACGCAG	886				join(15870..16043,16479..16484)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.5"
Db	874	TCTTAGACGCAG	886				join(15870..16043,16479..16484)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.5"
Db	874	TCTTAGACGCAG	886				/note="hypothetical ORF predicted by GENESH"
Db	874	TCTTAGACGCAG	886				this category is not included in IRGSP standard"
Db	874	TCTTAGACGCAG	886				complement(18932..25995)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.6"
Db	874	TCTTAGACGCAG	886				complement(join(18932..19648,19740..19836,20081..20194,20280..20444,20690..20827,20923..20999,21093..21207,21277..21382,21703..21764,21859..21993,22149..22331,22880..22936,23369..23417,23504..23652,24724..24771,25583..25995))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.6"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK120851"

Db	754	GGTAAACAGTGAAGCTGGTGGACCTTGCCCATGTGGCCGAGGTGATGGGTGATTGACCA	813				/organism="Oryza sativa (japonica cultivar-group)"
Qy	260	sAanPheLeuGlyGlyLeuCyseSerLeuIleLysPheValSerAapIleValProGluTh	280				/mol_type="genomic DNA"
Db	814	CAACTTCCTCGGGCGGCTCTGTAGCTGATCAAGTTTGTTCGTACATGTTCCAGAGAC	873				/cultivar="Nipponbare"
Qy	280	rProHisThrGln	284				/db_xref="taxon:39947"
Db	874	TCTTAGACGCAG	886				/chromosome="2"
Db	874	TCTTAGACGCAG	886				/clone="OSJNBa0047A17"
Db	874	TCTTAGACGCAG	886				join(2637..2780,3298..3360)
Db	874	TCTTAGACGCAG	886				join(2637..2780,3298..3360)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.1"
Db	874	TCTTAGACGCAG	886				/note="hypothetical ORF predicted by GENSCAN"
Db	874	TCTTAGACGCAG	886				this category is not included in IRGSP standard"
Db	874	TCTTAGACGCAG	886				complement(5864..9107)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				complement(join(5864..6755,8843..9107))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.2"
Db	874	TCTTAGACGCAG	886				/note="Rieske iron-sulfur protein contains EST(s): AU165272(C63256), AU068962(C51210) contains full-length cDNA(s): AK102815"
Db	874	TCTTAGACGCAG	886				/codon_start=1
Db	874	TCTTAGACGCAG	886				/product="putative ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial precursor"
Db	874	TCTTAGACGCAG	886				/protein_id="BAD26134.1"
Db	874	TCTTAGACGCAG	886				/db_xref="GI:49388912"
Db	874	TCTTAGACGCAG	886				/translation="MLRVAGRRLLTTALWRPAAAGARGPLAGSLPGDDEFRRPQR PRFVNDSPFFAASRGSTLVRNQDVLTELPAATVAVKPKPSAKIIVDEYNHRYQ PGDPSKRAFAYVLSSGGRFIVASLRLVLKFLVLSMSAKDLALASLEVLSSIEPG TTVTVKRGKVFIRRTTEEDINLANSVDIGSLRDPQDQDAERVKNPWLAVIGVCTHL GCILPNAAGDFGGWFCPCGSHYDISGRKRPAPYNLEVPYTSFLEENKLLIG"
Db	874	TCTTAGACGCAG	886				9904..10750
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				join(<9904..10067,10732..>10750)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				/note="start and end point are not identified"
Db	874	TCTTAGACGCAG	886				join(9904..10067,10732..10750)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.3"
Db	874	TCTTAGACGCAG	886				/note="predicted by GENSCAN etc."
Db	874	TCTTAGACGCAG	886				/codon_start=1
Db	874	TCTTAGACGCAG	886				/product="hypothetical protein"
Db	874	TCTTAGACGCAG	886				/protein_id="BAD26135.1"
Db	874	TCTTAGACGCAG	886				/db_xref="GI:49388913"
Db	874	TCTTAGACGCAG	886				/translation="MENATVTADLAPWSQQCRPVGEREEGGVAAAKKEGRQIQKE EKGGRRRSGSGIRKTM"
Db	874	TCTTAGACGCAG	886				complement(11599..12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				complement(<11599..>12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK109226"
Db	874	TCTTAGACGCAG	886				complement(11599..12180)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.4"
Db	874	TCTTAGACGCAG	886				/note="contains full-length cDNA(s): AK109226 non-coding transcript probably inactive due to including stop codon(s) in CDS"
Db	874	TCTTAGACGCAG	886				join(15870..16043,16479..16484)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.5"
Db	874	TCTTAGACGCAG	886				join(15870..16043,16479..16484)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.5"
Db	874	TCTTAGACGCAG	886				/note="hypothetical ORF predicted by FGENSEH this category is not included in IRGSP standard"
Db	874	TCTTAGACGCAG	886				complement(18932..125995)
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.6"
Db	874	TCTTAGACGCAG	886				complement(join(18932..19648,19740..19836,20081..20194,20280..20444,20690..20827,20923..20999,21093..21207,21277..21382,21703..21764,21859..21993,22149..22331,22880..22936,23369..23417,23504..23652,24724..24771,25582..25995))
Db	874	TCTTAGACGCAG	886				/gene="OSJNBa0047A17.6"
Db	874	TCTTAGACGCAG	886				/note="supported by full-length cDNA(s): AK120851"
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG	886				
Db	874	TCTTAGACGCAG					



ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS	1
TITLE	Sasaki, T., Matsumoto, T. and Yamamoto, K.
JOURNAL	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC clone: P0415B12
REFERENCE	Published Only in Database (2002)
AUTHORS	2 (bases 1 to 157419)
TITLE	Sasaki, T., Matsumoto, T. and Yamamoto, K.
JOURNAL	Direct Submission
COMMENT	Submitted (20-FEB-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan E-mail: tsasaka@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468 On Jun 28, 2004 this sequence version replaced gi:37497094. Genes were predicted from the integrated results of the following: GENSCAN ( <a href="http://CCR-081.mit.edu/GENSCAN.html">http://CCR-081.mit.edu/GENSCAN.html</a> ), FGENESH ( <a href="http://www.softberry.com/">http://www.softberry.com/</a> ), GeneMark.hmm ( <a href="http://opal.biology.gatech.edu/GeneMark/">http://opal.biology.gatech.edu/GeneMark/</a> ), GlimmerM ( <a href="http://www.tigr.org/tdb/glimmer/glmr_form.html">http://www.tigr.org/tdb/glimmer/glmr_form.html</a> ), RiceHMM ( <a href="http://rgp.dna.affrc.go.jp/RiceHMM/">http://rgp.dna.affrc.go.jp/RiceHMM/</a> ), SplicePredictor ( <a href="http://bioinformatics.laastate.edu/cgi-bin/sp.cgi">http://bioinformatics.laastate.edu/cgi-bin/sp.cgi</a> ), sim4 ( <a href="http://globin.cse.psu.edu/html/docs/sim4.html">http://globin.cse.psu.edu/html/docs/sim4.html</a> ), gap2 ( <a href="http://www.tigr.org/software/glimmer/">http://www.tigr.org/software/glimmer/</a> ), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr ( <a href="ftp://ncbi.nlm.nih.gov/blast/db">ftp://ncbi.nlm.nih.gov/blast/db</a> ) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to RGPSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0415B12 clone has an overlap with O3JNBa0047A17 clone (DDBJ: AP005749) at 5' end and an overlap with P0458B05 (DDBJ: AP004777) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <a href="http://rgp.dna.affrc.go.jp/Genomeseq.html">http://rgp.dna.affrc.go.jp/Genomeseq.html</a> .
FEATURES	Location/Qualifiers
Source	1. .157419 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosomes="2" /clone="P0415B12" complement (4313. .5807) /gene="P0415B12.1" complement {join(<4313. .4547,5568. .>5807)) /gene="P0415B12.1" /note="supported by full-length cDNA(s): AK121467" complement (4313. .5807) /gene="P0415B12.1" /note="contains full-length cDNA(s): AK121467 non-coding transcript probably inactive due to including stop codon(s) in CDS probably inactive due to too long 3'UTR in CDS" 6065. .7404 /gene="P0415B12.2"
Gene	complement (4313. .5807) /gene="P0415B12.1"
mRNA	complement {join(16881. .17069,17267. .17644,18400. .18594,18674. .18751,18816. .19001,19077. .19307,19386. .19739,19837. .20034,20118. .20288,20405. .20512,20611. .20829,20914. .21120,21227. .21445,21564. .21638,21998. .22166,22901. .23105)) /gene="P0415B12.4" /note="supported by full-length cDNA(s): AK065061" complement {join(16881. .17069,17267. .17644,18400. .18594,18674. .18751,18816. .19001,19077. .19307,19386. .19739,19837. .20034,20118. .20288,20405. .20512,20611. .20829,20914. .21120,21227. .21445,21564. .21638,21998. .22166,22901. .23105)) /gene="P0415B12.4" /note="contains EST(s): AU055859(S20105), AU085756 (C52347), C27605 (C52347), AU078722 (S20105) contains full-length cDNA(s): AK065061, AK063159" /codon_start=1 /product="RNA binding protein Rp120"
misc_feature	
Gene	

join(<6065. .6067,6850. .>7404)  
/gene="P0415B12.2"  
/note="start and end point are not identified"  
join(6065. .6067,6850. .7404)  
/gene="P0415B12.2"  
/note="predicted by FGENESH etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD25374.1"  
/db\_xref="GI:49388256"  
/translation="MNSRDQIPFLVYPRHEIVRARLAKEDGQHQQEQTVFEVKGDDGH  
SKNTLPSPASPAATPSALATIVASESLPEELQNHALLPTVKAQPLTSPKILSS  
SDLTSSVTITTAANSEKTAAMVAATVCSMKGLEKPTNPRTVREHVVGSRGLCR  
TVSGQVGEVLGEGSCQITVPCH"  
complement (8669. .15038)  
/gene="P0415B12.3"  
complement {join(8669. .9060,9161. .9334,9438. .9569,  
9640. .9738,9929. .10000,10080. .10136,10206. .10379,  
10460. .10554,10727. .10826,10921. .11073,11145. .12275,  
12870. .12980,13078. .13240,13613. .13707,13793. .14053,  
14178. .14331,14775. .15038))  
/gene="P0415B12.3"  
/note="supported by full-length cDNA(s): AK121782"  
complement {join(8929. .9060,9161. .9334,9438. .9569,  
9640. .9738,9929. .10000,10080. .10136,10206. .10379,  
10460. .10554,10727. .10826,10921. .11073,11145. .12275,  
12870. .12980,13078. .13240,13613. .13707,13793. .14053,  
14178. .14331))  
/gene="P0415B12.3"  
/note="contains EST(s): AU082722 (R4038), AU032417 (R4038)  
contains full-length cDNA(s): AK121782"  
/codon\_start=1  
/product="gamma-tubulin complex component 5-like"  
/protein\_id="BAD25375.1"  
/db\_xref="GI:49388257"  
/translation="MVLQNLNAGLMEADSGASFIRKLQLSVSDGLPAAVPVELSTQE  
HELKSVPEVLQGPDTLLYNDKTAPGCEKAGIVVSHLSQTSRAVLKPPFAATQCL  
KQVELFVGRVRSCHGTITLSAFASSVDSWMLRKAALKEEQFLSVERITILGL  
TDSMSLCGAELHYQVVGAVPDAFNSGQMASSEVAHVAVHLLFKLNBSVCLVED  
GEGEYHMLLVIFAGTLPLPYQCLDSWLYDGLDDPNEEMFVANKAVITDQPAWEM  
SYMLRVGRPTDSSLTADNESIRKKELINQETTAALAKLSSNQGCADILCPVLKD  
IARALISAGKSPQLVQHVEITHRIQTRVIEHFNIDQHNYSIQKFRPDTSSIRIQD  
KREDIIEBTSQFGNNACMDFLTLSSEFLICLSGLENGDHVDVLRKLCADNAPVN  
KTIHVKSNVOETEVCCENSEKTLKLRDATSGRDYDGMKTLKNAVMDPTFV  
PGDHQDVSTVESYFNLSVCENPGITACQEMLEKNKNSWDLNLSKSFHLPPLNDEN  
IRKSIIFGRDSSGIPDTSTTYPPLDGTDFKGFQDDSEYIRQEDDRKTLALY  
TFPTLLPCNVENNPISLEILPLQKSTLASRAKFTQSMSLRDPVGLIQCLSKICI  
KRQVDHLKQILSKLMDGWRMLDELFLVLRAYLGLSDMLQQLFVITFDLKGPNWD  
DDFELNTLQESIRNSADKMLLTAPDSLVSLSLAKHDTNRDETTSIRKRAQGFIE  
ALDLVNTFYKYSVPLDLIVNTEALKYQNVMAFLVKVRAKFTILDETETKMMWKGGSST  
MHNFKQLHIVQKLLHFVDFAHQYVMDRVYHSANTELCDGNASATILDEVHVEAYL  
SSIQRQCFVADKUALIASRVKTLILGALDFHNIETGLTGCTGTPAPAKARCEMEVDR  
TEKQFCECVFLIRILSPKLNHPPHPLADLVTRINYNHYNMSDSGSFSAIPGSRPR"  
complement (16514. .23105)  
/gene="P0415B12.4"  
complement {join(16514. .17069,17267. .17644,18400. .18594,  
18674. .18751,18816. .19001,19077. .19307,19386. .19739,  
19837. .20034,20118. .20288,20405. .20512,20611. .20829,  
20914. .21120,21227. .21445,21564. .21638,21998. .22166,  
22901. .23105))  
/gene="P0415B12.4"  
/note="supported by full-length cDNA(s): AK065061"  
complement {join(16881. .17069,17267. .17644,18400. .18594,  
18674. .18751,18816. .19001,19077. .19307,19386. .19739,  
19837. .20034,20118. .20288,20405. .20512,20611. .20829,  
20914. .21120,21227. .21445,21564. .21638,21998. .22166,22901. .23105))  
/gene="P0415B12.4"  
/note="contains EST(s):  
AU055859(S20105), AU085756 (C52347), C27605 (C52347), AU078722 (S20105)  
contains full-length cDNA(s): AK065061, AK063159"  
/codon\_start=1  
/product="RNA binding protein Rp120"



Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
Japanese rice  
Science 301 (5631), 376-379 (2003)

## JOURNAL MEDLINE

22752273

12869764

2 (bases 1 to 1570)

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arahawa, T., Carninci, P., Doi, K.,  
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, J., Imamura, K.,  
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
Kanguchi, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
Masuda, H., Matubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
Numasaki, R., Oneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
Yamada, H., Yamamoto, M., Yasunishiki, A., Yazaki, J., Yokomizo, S., and  
Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001)

Shoichi Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,

Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica

rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K.,

Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,

Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and

Yamamoto, M.

FAIS Genome Sequencing &amp; Analysis Group: Ootomo, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,

Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,

Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,

Yoshimura, A., Matubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center

and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arahawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,

Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,

Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,

Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,

Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. .1570

/organism="Oryza sativa (japonica cultivar-group)"

/mol type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="J023018G11"

## FEATURES

source

## ORIGIN

Alignment Scores:  
Pred. No.: 2,966-77 Length: 1570  
Score: 1144.50 Matches: 217  
Percent Similarity: 84.10% Conservative: 21  
Best Local Similarity: 76.68% Mismatches: 40  
Query Match: 75.25% Indels: 5  
DB: 8 Gaps: 3

US-10-042-894A-8 (1-289) x AK072296 (1-1570)

QY 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeu 21  
Db 384 TCCGACCTTGGCGCCCGCGAGCACAGGTGGGGGGGACCGCGCTCCGCCACAGCTG 443  
QY 22 GlyProLeuIleAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArgGly 41  
Db 444 GCGCGCGCTTGGCGCGCGAGGGGCTTCTTCAAGACCCCTCCAGACCGGGGAGCGCGG 503  
QY 42 GluHisGluValAlaPheTyrlsGluAlaPheSerAlaHisAlaAlaValProAlaArgIle 61  
Db 504 GAGCAGAGGGCGCTTCTTACCGCGCTTACCGCGCGACCGCGCTCCGCCCGGGTTC 563  
QY 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81  
Db 564 CGGGCGCGCTTCTTCCCGCGCTTCCAGCGCACCGCGCTTCTCCCGCGCCCGACGACG 623  
QY 82 ---GlyGluProHisProHisGlnValLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 624 GCGCGCGCGCTTACCGCGACATCTCTCTCGAGACCTCTCTCGCGCGCTCCCGTCCCC 683  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
Db 684 TCGGTTCGCGCGCTTCAAGTCCGCGCGCTTCCAGTTCGCGCGCGCGATCCCGGACCCCTAC 743  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 744 GTCCGCAAGTGCTCGCCAAAGACCGCGAGACCGCGCGCTCTCTCGCGTTCGCGCTC 803  
QY 141 SerGlyValArgValValGlyProGlu---GlyAlaValTrpArgThrGluArgProGlu 159  
Db 804 TCGCGCGTTCGCGGTGGTTCGATGCCGCGCGCGCGCGTGTGGCGCGCGCGCGGTTCGAG 863  
QY 160 ValLysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerVal 179  
Db 864 CTGAGAGGATCGACCGCGCGCGGTTCGCGCGCTCTCCGCGCTTCTCCACGGCG 923  
QY 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyVal 199  
Db 924 GCGCGCGCGCTTGGACTTCGCGCTTCGCGCGCGGTGTACGAGGGGAGCGCGCGCTC 983  
QY 200 LeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyr 219  
Db 984 CTGGTTCAGTTCGCGGAGCTCAAGCGGTGTTCGAGGAGCAAAACCTCTTACCATTTCTAC 1043  
QY 220 SerAlaSerIleLeuLeuGlyTyrAsp-----AlaAlaAlaValAlaAlaGlyGly 236  
Db 1044 TCGCGGTTCGATCTGTTCGCTACGACCATTCGCGCGCGCGCGCTCTCCCGAGGT 1103  
QY 237 AspGlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256  
Db 1104 GGAAGCGCGGTGTGAAGGTGAAGGTGTGGTGTGGTGTTCGCGCATGTTCGACGAGCGG 1163  
QY 257 ValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276  
Db 1164 GTGATTCACCAACTTCTTGGCGCGGTCTCTCTCTCATCAAGTTCTCGCGACATT 1223  
QY 277 ValProGlu 279  
Db 1224 GTCCGCGAG 1232

AX513580 643 bp DNA linear PAT 05-OCT-2002

RESULT 10  
AX513580  
LOCUS

DEFINITION Sequence 17 from Patent WO02059324.  
ACCESSION AX513580  
VERSION AX513580.1 GI:23559681  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1  
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalaki, J.A. and Cahoon, R.E.  
TITLE Novel inositol polyphosphate kinase genes and uses thereof  
JOURNAL Patent: WO 02059324-A 17 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. .643  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,456-52 Length: 643  
Score: 808.50 Matches: 162  
Percent Similarity: 93.14% Conservative: 1  
Best Local Similarity: 92.57% Mismatches: 10  
Query Match: 53.16% Indels: 3  
DB: 6 Gaps: 0  
US-10-042-894A-8 (1-289) x AX513580 (1-643)  
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisLys 20  
DB 117 ATCCCGACGCTCCACCGCGCGGAGCACCAGTCCGCGCTCCACCGCGCTCCGCGCAAG 176  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 177 CTGGGCGCGCTCATCGACGCTCCGCGCTCTCTACAGCGCTCCAGCGCGGCGCGT 236  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 237 GGGGAGCAGGAGTCCGCTCTATGAGCGCTTCTCGCCGACCGCGCGCGCGCGC 296  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 297 ATCCGAGACACCTTCTTCCCGCGGTTCCACGCGACGCGACTCTCCCGCGCGCGCAG 356  
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
DB 357 CCCGGGAGCGGATCCGACCTCTGCTCGACGCTCTCGCGCGCGCGCGCGCGC 416  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
DB 417 TGGCTCGCAGACATCAAGATCGCGCGCATCATCGTGGCCACCGAGTTCGCGGAGCGCTAC 476  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 477 ATCCGCAAGTACCTGNGCCAGGACCGCGGACCGCGCGCTTCTGCTCGGATTCGCGTC 536  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArg-ProGluVal 160  
DB 537 TTGC--GTCCGAGTCTGCGCGCGCGCGCGCGCTGTGCGGACGCGCGCGCGG 594  
QY 160 lLysAlaMetAspThr-AlaGlyValArgValLeuArg 173  
DB 595 GAANGCTATGACACCGCTCGGNGCGCGCGGCTGCTTCG 635  
RESULT 11  
AX513576  
LOCUS Sequence 13 from Patent WO02059324. linear PAT 05-OCT-2002  
DEFINITION  
ACCESSION AX513576  
VERSION AX513576.1 GI:23559677  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1  
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalaki, J.A. and Cahoon, R.E.  
TITLE Novel inositol polyphosphate kinase genes and uses thereof  
JOURNAL Patent: WO 02059324-A 17 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. .1020  
/organism="Parthenium argentatum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:35935"  
21. .908  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52957.1"  
/db\_xref="GI:23559678"  
/translation="MLKADPQVAGHEAGLGLPLIDDSGRFYKPLQGNRSGSEVA  
FYSPSSNNNIPEHIRKFFIYGTIKMASTGSDHPHVLQDLTSAHVPNSYMDKI  
GSRWAPAESEAYIAKCLKKDRSTIPLGFRISGLQVYIDGSGFYKPHRNPYMRKT  
PADVRLLLKRFVSSNPSEMEWRTGLGPCSLASFYVGGPNGILAOLMELKLTWFDQT  
IYHFACSFPIFEKRLVLKGARSNAEVLIDPAHVTDGNGVIDHNFLOGGLCSLIRFI  
SDILSETKDCNGTNGQVEL"  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,056-47 Length: 1020  
Score: 742.50 Matches: 149  
Percent Similarity: 67.83% Conservative: 45  
Best Local Similarity: 52.10% Mismatches: 79  
Query Match: 48.82% Indels: 13  
DB: 6 Gaps: 5  
US-10-042-894A-8 (1-289) x AX513576 (1-1020)  
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisLysLeuGlyPro 23  
DB 24 CTCAGGCGCCAGATCATCAGGTTCGTGACATGAAGCTGGGCTCGGGAAGCTGGCCCA 83  
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43  
DB 84 CTCATTGATGATTTCAGGCGCGTTCACAAACCACTGCGGGTGATAACCGTGGGTACAG 143  
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgGlyArgAsp 63  
DB 144 GAAGTAGCCTTTATGAATCATTTCTTACCAATAATATTCAGAACACATACGC-- 200  
QY 64 ThrPheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83  
DB 201 AAATCTTCTTCTATATATTATGACCAACAAATCATG-----AAGGCATCCATGGCTCT 254  
QY 84 ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103  
DB 255 GACCATCTCCATCATGTTTCAAGATCTTACATCATGCTCATGCAACCCATCTGTAATG 314  
QY 104 AspIleValIleGlyAlaIleThrTrpProSerSerProGluProTyrIleAlaLys 123  
DB 315 GACATCAAAATCGGCTCCAGAACATCGGCGCCAGAAAGCTTCCAGGGGTACATTGCAAAA 374  
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143  
DB 375 TGCCTTAAAAAGGATAGGGAACACACAGATATTCCATTGGGATTCAGGATCTCCGGGCTG 434  
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162  
DB 435 CAAGTCTATATCATGATGCGTTCAGGCTTTTATAAGCCTCATAGAAATATACGCGTAAA 494  
QY 163 MetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerValAlaAspGlu 182



Db	495	ACCGGCCGACGCTGATGTTAGACTACTTCTTAGAAATTTGTTCTTAACCCCGTCTGCA	554
Qy	183	-----GlyMetAspCysAlaLeuAlaAlaAlaValTyrGly	194
Db	555	GAGATCGAAATCGGCACAGCGCTAGCCCGGATTTGTTTGTAGCATCTTTGTTATGGT	614
Qy	195	GlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTppPheGluGluGlnThr	214
Db	615	GGCCCTAATGGGATATTAGCTCAACTGATGGAATTGAAGACATGGTTGAAGATCAACA	674
Qy	215	LeuPheHisPheTyrSerAlaSerLeuLeuLeuGlyTyrAspAlaAlaAlaValAlaAla	234
Db	675	ATTTACCACCTTCTATGCTTGTCTTTTGTGTTTGTATCTTTGTAAGAGAGGTTGGTCTTAA	734
Qy	235	GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly	254
Db	735	GGTGCT---CGGTCAACCGCAGAGTCAAACTATTGATTTTGTCTTGTATGATTTCTT	791
Qy	255	AspGlyValIleAspHisAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSer	274
Db	792	AATGGTGTTATTGATCACAATTTCTTGGTGCGCTCTGTTCTTTGATAAAGTTCATTCT	851
Qy	275	AspIleValProGluThr	280
Db	852	GACATACCTTCGGGAGCA	869
RESULT 12			
AX513574		1195 bp	DNA
LOCUS		Sequence 11 from Patent WO02059324.	linear
DEFINITION			PAT 05-OCT-2002
ACCESSION		AX513574	
VERSION		AX513574.1	GI:23559674
KEYWORDS			
SOURCE		Eucalyptus grandis	
ORGANISM		Eucalyptus grandis	
REFERENCE		Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.	
AUTHORS		Novel inositol polyphosphate kinase genes and uses thereof	
TITLE		Patent: WO 02059324-A 11 01-AUG-2002;	
JOURNAL		PIONEER HI-BRED INTERNATIONAL, INC. (US)	
FEATURES		Location/Qualifiers	
source		1..1195	
		/organism="Eucalyptus grandis"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:71139"	
CDS		116..1048	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD52956.1"	
		/db_xref="GI:23559675"	
		/translation="MLKVPDPHOVAGHRGDGKGLPLVDDSGRFYKPLQSDHRGDEVA	
		FYESFYNTPEIPGHRKFPAPFKTTEASDGGSPDPLVLELDVSGRTNPSLMDIK	
		TGSRWYPSASEYIQKLEKDRNSTSVSLGFRISGLRVYQNSAGWQPEKKVYVF	
		NADGVAURKRVNSLSGPNVDPLVASKYVCHRGRIQLAQLEKFWFVOTNTH	
		FYCSLIIDYRESALDGAHPKVLVDFAHMDGHVGDHNFILGGLCSVIKFRIDIA	
		DEDNKCAKCEVNLGLKENGIFYKSTPEPLDHEAC"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1.58e-45	Length:	1195
Score:	722.50	Matches:	149
Percent Similarity:	66.31%	Conservative:	38
Best Local Similarity:	52.84%	Mismatches:	84
Query Match:	47.50%	Indels:	11
DB:	6	Gaps:	6
US-10-042-894A-8 (1-289) x AX513574 (1-1195)			
Qy	4	LeuHisProGluHisGlnValAlaGlyHisAenPheAlaSerLeuLysGlyPro	23



AUTHORS Dodds, P.N., Lawrence, G.J. and Ellis, J.G.  
TITLE Contrasting modes of evolution acting on the complex N locus for rust resistance in flax  
JOURNAL Plant J. 27 (5), 439-453 (2001)  
MEDLINE 21461288  
PUBMED 11576428  
REFERENCE 2 (bases 1 to 25054)  
AUTHORS Dodds, P.N.  
TITLE Direct Submission  
JOURNAL Submitted (22-WAR-2001) Dodds P.N., Plant Industry, CSIRO, GPO Box 1600, Canberra, ACT 2601, Australia  
FEATURES  
source  
1. .25054  
/organism="Linum usitatissimum"  
/mol\_type="genomic DNA"  
/variety="Bombay"  
/db\_xref="taxon:4006"  
/haplotype="N"  
join(2881..3431,3533..4652,4770..5045,5140..6519)  
/genes="Ngc-D"  
join(2881..3431,3533..4652,4770..5045,5140..6519)  
/genes="Ngc-D"  
/notes="N rust resistance gene: TIR-NBS-LRR protein"  
/codon\_start=1  
/product="Ngc-D protein"  
/protein\_id="CAC35321.1"  
/db\_xref="GI:13509207"  
/db\_xref="GOA:Q9ARC4"  
/db\_xref="UniProt/TREMBL:Q9ARC4"  
/translation="MMRSNGSTDSFHCSSADLTFFPLPSGEYEVFLSFRGPDVVRK  
TFADHLYTSLVRKPTFDEBELKGGTIGPSLIRAITESKIYIPILTQNYASSKWC  
LQELAKWVCNGGAKGQHIIIPVYFMDPRDVRHPSGPKYKEAFEGHNSKHDPET  
VLEWKGALQEVGMKGHISELTGQAVVDKI FTEVELHRLANYATLDELVGTIDFSV  
DEMVLNLDSSTSEKIGIYGMGRGKTTATAVYKNSMOPERCCTFLNIRETLTKN  
DGVVALQKVISDIILKRCQAKNASDGVQMIERVSRRHKIFVLDVNVESFPDDIF  
GKLTAFSDFRFLVTRDARTLERLGRCKLFKHEGNSHDHSLFKFARHAPGVDPYPPD  
YASLCEQFVGSGLPALKVTGSLLFRTKSFMDCKLIELKAI PAVNVQYRLKISYN  
ELTDNEKQIFLDVACLFGAKKEIPIYMSDCGFTPTTIRTLQVSLRVINDNEEFM  
MDHRLDGLRAIVCESONLYKRSIWNNDADI LKRNENDCVCLARVDMRGEFGA  
LTNEEPKQSRRLRELVNGDSGNFKNVLPISLRWLRVVHGDP CSPGLNKLKMLIEL  
EVSVDTSWEGWNEIKAAKGLKVHLMCKGKLEKVPDLSTCRGLELLRFSICRMHGE  
LDIRPKDKLDIDIFQTRITAKLVEVESLQNLQDLVSGSLIEFAGISKLASLEYL  
NRTLNKHDEVLPNGELKLLISFSSALPSLFLDRLVRYSTLRELAVRCLILEKL  
RLNREVEGHI GPGLGELKLCFLRDLAPNLDNLGLENLVLKELAVRCLILEKL  
PISLAELTKHLVIGOCNLTGELIYGLANGLSHLEISGCPCLTVVSLVSLNLTGL  
LHLESLERLVNGSSIRKLDLTGLVKLEILQPDSCQTQLTEIRGLGLESIRQLHMSRC  
QIKSLPNLSUGKILSIYILEKCRHLKVEYRLEELKRLDFNTHRG"  
17600..18457  
/note="ORF"  
/function="unknown"  
related to human inositol hexakisphosphate kinase 2  
protein"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAC35324.1"  
/db\_xref="GI:13509210"  
/db\_xref="GOA:Q9ARC1"  
/db\_xref="UniProt/TREMBL:Q9ARC1"  
/translation="MLKVEPHQVAGHQDINGLLGPLVDDSGRFYKPKLQDGERGATENA  
FTYTSSTNKTIPQVRRFFAQGTQSIASDGSGLHPLHILEDLTSTRHLVCMDIK  
IGSRWYPEASQAYIEKCLKDQVSSNPLGFRIISGLVYNGKESSEVVLKPERKLL  
LNLNADDEVRLVKRFVSSNLKQDQPCFAAVVSSGNGILAQLLEKAWFEDQTIYH  
FNCSVLMLYEKKTWNGGSESLGAAVKLIDFAHVTGNGVIDHNFGLGCLSLIK  
FISENQIML"  
join(20265..20815,21104..22202,22415..22690,22818..24218)  
/genes="Ngc-B"  
join(20265..20815,21104..22202,22415..22690,22818..24218)  
/genes="Ngc-B"  
/note="N rust resistance gene homolog: TIR-NBS-LRR  
protein"  
/codon\_start=1  
/product="Ngc-B protein"  
/protein\_id="CAC35325.1"  
/db\_xref="GI:13509211"  
/db\_xref="GOA:Q9AR40"  
/db\_xref="UniProt/TREMBL:Q9AR40"  
/translation="MMRSNTDSIGSFLSRSSVDPTLLPLPSGEYEVFLSFRGPDVVRK  
TFADHLYAWLVRSKIRTFDEBLQGETIGSSLIQAITESKIYIPILTQNYASSKWC  
LQELAKWVCWNGGKGQHIIIPVYFMDPRDVRHPSGPKYKEAFEGHNSKHDPET

AUTHORS Dodds, P.N., Lawrence, G.J. and Ellis, J.G.  
TITLE Contrasting modes of evolution acting on the complex N locus for rust resistance in flax  
JOURNAL Plant J. 27 (5), 439-453 (2001)  
MEDLINE 21461288  
PUBMED 11576428  
REFERENCE 2 (bases 1 to 25054)  
AUTHORS Dodds, P.N.  
TITLE Direct Submission  
JOURNAL Submitted (22-WAR-2001) Dodds P.N., Plant Industry, CSIRO, GPO Box 1600, Canberra, ACT 2601, Australia  
FEATURES  
source  
1. .25054  
/organism="Linum usitatissimum"  
/mol\_type="genomic DNA"  
/variety="Bombay"  
/db\_xref="taxon:4006"  
/haplotype="N"  
join(2881..3431,3533..4652,4770..5045,5140..6519)  
/genes="Ngc-D"  
join(2881..3431,3533..4652,4770..5045,5140..6519)  
/genes="Ngc-D"  
/notes="N rust resistance gene: TIR-NBS-LRR protein"  
/codon\_start=1  
/product="Ngc-D protein"  
/protein\_id="CAC35321.1"  
/db\_xref="GI:13509207"  
/db\_xref="GOA:Q9ARC4"  
/db\_xref="UniProt/TREMBL:Q9ARC4"  
/translation="MMRSNGSTDSFHCSSADLTFFPLPSGEYEVFLSFRGPDVVRK  
TFADHLYTSLVRKPTFDEBELKGGTIGPSLIRAITESKIYIPILTQNYASSKWC  
LQELAKWVCNGGAKGQHIIIPVYFMDPRDVRHPSGPKYKEAFEGHNSKHDPET  
VLEWKGALQEVGMKGHISELTGQAVVDKI FTEVELHRLANYATLDELVGTIDFSV  
DEMVLNLDSSTSEKIGIYGMGRGKTTATAVYKNSMOPERCCTFLNIRETLTKN  
DGVVALQKVISDIILKRCQAKNASDGVQMIERVSRRHKIFVLDVNVESFPDDIF  
GKLTAFSDFRFLVTRDARTLERLGRCKLFKHEGNSHDHSLFKFARHAPGVDPYPPD  
YASLCEQFVGSGLPALKVTGSLLFRTKSFMDCKLIELKAI PAVNVQYRLKISYN  
ELTDNEKQIFLDVACLFGAKKEIPIYMSDCGFTPTTIRTLQVSLRVINDNEEFM  
MDHRLDGLRAIVCESONLYKRSIWNNDADI LKRNENDCVCLARVDMRGEFGA  
LTNEEPKQSRRLRELVNGDSGNFKNVLPISLRWLRVVHGDP CSPGLNKLKMLIEL  
EVSVDTSWEGWNEIKAAKGLKVHLMCKGKLEKVPDLSTCRGLELLRFSICRMHGE  
LDIRPKDKLDIDIFQTRITAKLVEVESLQNLQDLVSGSLIEFAGISKLASLEYL  
NRTLNKHDEVLPNGELKLLISFSSALPSLFLDRLVRYSTLRELAVRCLILEKL  
RLNREVEGHI GPGLGELKLCFLRDLAPNLDNLGLENLVLKELAVRCLILEKL  
PISLAELTKHLVIGOCNLTGELIYGLANGLSHLEISGCPCLTVVSLVSLNLTGL  
LHLESLERLVNGSSIRKLDLTGLVKLEILQPDSCQTQLTEIRGLGLESIRQLHMSRC  
QIKSLPNLSUGKILSIYILEKCRHLKVEYRLEELKRLDFNTHRG"  
4553  
/genes="Ngc-D"  
/standard\_name="n-X50"  
/notes="Site of 11 bp insertion in n-X50 rust susceptible  
mutant allele"  
/evidence=experimental  
5608  
/genes="Ngc-D"  
/standard\_name="n-X128A"  
/notes="Site of LUTE transposon insertion in n-X128A rust  
susceptible mutant allele"  
/evidence=experimental  
6397..6629  
/standard\_name="n-X139"  
/notes="233 bp deletion in n-X139 rust susceptible mutant  
allele"  
/evidence=experimental  
9613..10515  
/function="unknown"  
/notes="ORF  
related to human inositol hexakisphosphate kinase 2  
protein"  
/codon\_start=1  
/product="hypothetical protein"



Qy	4	LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro	23
Db	4	CTCAAGGTCCTGAAACACCAAGTTGCTGTGTCACATTCGTAGTATGGAGAGCTCGGTCCA	63
Qy	24	LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis	43
Db	64	CTCGTAGATGACCAAGCGCGTCTTCAAGCCACTTCAGGGAGATTCTCTGTGCGGAACAC	123
Qy	44	GluValAlaPheTyrGluAlaPheSerAlaHisAlaAalaValProAlaArgIleArgAsp	63
Db	124	GAGGCTAAGTTCATGAGTCTTTCACATCGAACATGAAGGTTCCAGATCACATCCAT	180
Qy	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu	83
Db	181	AGTACTTCCCGGTGATACCGCACTCAGCTAGTT-----GAAGCATCTGATGGATCT	234
Qy	84	ProHis---ProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal	102
Db	235	GGCAAGCTTCTCATCTGTTCTGTGATGATGTTGTTTCAGGGTACGCAACCGTCGGTA	294
Qy	103	AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla	122
Db	295	ATGGATGTTAAGATTGGATCTAGACATGTCACCGGATGTATCAGAAGAATTACTTCAAG	354
Qy	123	LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly	142
Db	355	AAATGTTAAGAAGATAGACAGACACCCAGGTTTCGTGGGGTTTCAGGTTTCAGGT	414
Qy	143	ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys	161
Db	415	TTTAAAGATTTTTCATCACCAAGAATCAAGTTTTTGGAGAGCTCAGAAGAAGCTTGTCTT	474
Qy	162	AlaMetAspThrAlaGlyValArgValLeuArgTyrValSer-----SerVal	179
Db	475	GGGTATAATGCAGATGCTGCTAGATTGGCTCTCAGGAAGTTTGTGTCATCGAACTCTCC	534
Qy	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly	197
Db	535	GCTGACTCTAATCTTGACACCAAACTGTGCTTTTGGCATCAGAGGTTTATGGCGTTGTAAC	594
Qy	198	GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHis	217
Db	595	GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGGTTTCGAAACCCCAACGCTTTACCAT	654
Qy	218	PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly	236
Db	655	TTCAATTCCTGCTCGATTCTGATGATTTACGAGATGAATCAATCTTGTGCAAGGAGGA	714
Qy	237	AspGlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAlaGlu	253
Db	715	GATGATGACCGGCACACCGGCACAAAGTGAAGCTGGTGGATTTCCGCGATGTTCTGTAT	774
Qy	254	GlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal	273
Db	775	GGAACCGGTGTCATCGACCATATTTCTCTGGTGGACTCTGCTCTTTTCATAAAGTTTCATC	834
Qy	274	SerAspIleVal 277	
Db	835	AAAGATATTCTT 846	
RESULT 15			
AY072621			
LOCUS	AY072621	903 bp	mRNA linear PLN 20-JAN-2002
DEFINITION	Arabidopsis thaliana AT5G61760/mac9_60 mRNA, complete cds.		
ACCESSION	AY072621		
VERSION	AY072621.1	GI:18252262	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 903)		

AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.			
TITLE	Arabidopsis ORF clones			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 903)			
AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.C., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-JAN-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA			
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.			
	The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.			
	Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.			
FEATURES	Location/Qualifiers			
source	1..903			
	/organism="Arabidopsis thaliana"			
	/mol_type="mRNA"			
	/db_xref="taxon:3702"			
	/chromosome="5"			
	/clone="U13510"			
	/ecotype="Columbia"			
	/note="This clone is in pUNI 51"			
CDS	1..903			
	/note="unknown protein"			
	/codon_start=1			
	/product="AT5G61760/mac9_60"			
	/protein_id="AAL62012.1"			
	/db_xref="GI:18252263"			
	/translation="MLKVEHQVAGHISDGKGLPLVDQGRPFKPLQDGRGEHEAK FYESFTSNMKVPDHIHRYFPVYHGTQLVASDGSGLPHLVDVVGSTANPSVMDVK IGRTWYPDVSEYFKKIKQDRTTSLGFRVSGFKIPDQESSFWABKRLVLYG NADGARLARLKFVSSNPADSLNTPNCAPASEYVGGNGILAQLELKMWFQTLYH FNSCSILMTYENESILMQGDDAPAPRAQVKLVDFARVLVDNGVVIDHNFLLGGLCSFTK FIKDILQSVKEHDETDTSLENGR"			
ORIGIN				
Alignment Scores:				
Pred. No.:	4..57e-45	Length:	903	
Score:	714.50	Matches:	150	
Percent Similarity:	66.55%	Conservative:	39	
Best Local Similarity:	52.82%	Mismatches:	82	
Query Match:	46.98%	Indels:	13	
DB:	8	Gaps:	8	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 21:48:51 ; Search time 5916 Seconds  
(without alignments)  
11008.082 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgagtcgctgcac.....ataaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1344	6	AX513570 Sequence
2	865.4	64.4	923	6	AX513568 Sequence
3	844.8	62.9	1169	6	AX513564 Sequence
4	837.4	62.3	923	6	AX513566 Sequence
5	740.4	55.1	3416	6	AX513583 Sequence
6	686	51.0	899	6	AX513578 Sequence
7	546.4	40.7	149142	8	AP005749
8	546.4	40.7	157419	8	AP004772
9	544.8	40.5	1570	8	AX072296
10	470.4	35.0	643	6	AX513580
11	298.4	22.2	519	6	AX513581
12	191.6	14.3	353	6	AX513582
13	184.6	13.7	1195	6	AX513574
14	181.6	13.5	25054	8	LUS310150
15	179	13.3	1105	6	AX513572
16	149.8	11.1	1020	6	AX513576
17	148.6	11.1	1390	8	AY136378
18	148.6	11.1	116763	8	ATT211
19	148.2	11.0	905	8	BT000196

20	147.8	11.0	115040	8	AC124954
21	146.6	10.9	858	8	BT011900
22	146.6	10.9	1243	8	AY087217
23	145.8	10.8	1068	8	ATH404678
24	145.8	10.8	130078	2	AC127018
25	145.6	10.8	903	8	ATH243592
26	145.6	10.8	903	8	AY072621
27	145.6	10.8	903	8	AY147936
28	145.6	10.8	1166	8	AF412073
29	145.6	10.8	1170	8	ATH245521
30	145.6	10.8	57246	8	AB010069
31	145.4	10.8	861	8	AY147935
32	143.6	10.7	1130	8	AY085862
33	137.6	10.2	84869	12	AY350714
34	83.8	6.2	43428	8	AP004981
35	75.2	5.6	110000	2	LMFLCHR32_06
36	75	5.6	110000	2	LMFLCHR36_03
37	74.8	5.6	296500	1	SCO939128
38	73.8	5.5	125020	9	AF429315
39	73.6	5.5	110000	2	LMFLCHR36_31
40	73.4	5.5	300800	1	SCO939112
41	72.4	5.4	113193	1	AF357202
42	72.4	5.4	113193	6	AX703543
43	71.4	5.3	135638	1	AF484556
44	70.4	5.2	224	8	AY022706
45	70.4	5.2	78220	2	AC023212

#### ALIGNMENTS

RESULT 1	AX513570	Sequence 7 from Patent WO02059324.	1344 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	AX513570	Sequence 7 from Patent WO02059324.				
DEFINITION	AX513570	Sequence 7 from Patent WO02059324.				
ACCESSION	AX513570	Sequence 7 from Patent WO02059324.				
VERSION	AX513570.1	GI:23559670				
KEYWORDS						
SOURCE	Zea mays					
ORGANISM	Zea mays					

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

#### REFERENCE

- 1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.  
Novel inositol polyphosphate kinase genes and uses thereof  
Patent: WO 02059324-A 7 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)

#### FEATURES

source

1. 1344  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
52..921

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD52954.1"

/db\_xref="GI:23559671"

/translation="MSDLHPPEHQVAGHRASAKGLPLIDSGSLFYKPLQADGRGEHE  
VAFYFAFSAHAAPARIDTFFPHGHTRLLPTEAQGPGEHPHVLIDLLAGFOAPCV  
ADIKIGAITWPPSPPEPIYAKLAKDRGTTSLVIGFRVGVVVEGAVWRTEPEV  
KAMDTAGRVRLRRYVSSVADGMDCAAAYGKGVLSQLRELKAWFEQTLTFHF  
YMSILLYGDAVAAGDGGVTVKLVDFAHVAEGDGVIDHNFGLGLCSLIKPFVSDI  
VPETPHTQLGPS"

#### ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 6.8e-201;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGTCAGTCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGAC 60  
|||  
Db 1 GCACGAGTCAGTCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGAC 60  
|||

```
Qy 61 CTCACCGCGGAGCACCAGTCCGCGGCACCGCGCTCCGCGCAGCAGCTGGGCGG 120
Db 61 CTCACCGCGGAGCACCAGTCCGCGGCACCGCGCTCCGCGCAGCAGCTGGGCGG 120
Qy 121 CTATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGCGCGGAGCCGCTGGGGAGCAC 180
Db 121 CTATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGCGCGGAGCCGCTGGGGAGCAC 180
Qy 181 GAGTCGCTCTATAGGGGTTCTCCGCCACCGCGCGCTCCGGCGCGGATCCGAGAC 240
Db 181 GAGTCGCTCTATAGGGGTTCTCCGCCACCGCGCGCTCCGGCGCGGATCCGAGAC 240
Qy 241 ACCTTCTTCCCGCGGTTCCAGGACCGGACTCTCCCAACCGAGGCGGAGCCGGGAG 300
Db 241 ACCTTCTTCCCGCGGTTCCAGGACCGGACTCTCCCAACCGAGGCGGAGCCGGGAG 300
Qy 301 CCGCATCTCACTCTGCTCGACGACTCTCCGCGGGTTTCAGGGCGCTCGGTCGCA 360
Db 301 CCGCATCTCACTCTGCTCGACGACTCTCCGCGGGTTTCAGGGCGCTCGGTCGCA 360
Qy 361 GACATCAAGATCGCGGCATCACTGCGCACCGGATTCGCGGAGCCCTACATCGCCAAG 420
Db 361 GACATCAAGATCGCGGCATCACTGCGCACCGGATTCGCGGAGCCCTACATCGCCAAG 420
Qy 421 TGCCTCGCAGGACCGGGACCAAGCGGTTCTCGGATTCGCGGTCGCGGCTC 480
Db 421 TGCCTCGCAGGACCGGGACCAAGCGGTTCTCGGATTCGCGGTCGCGGCTC 480
Qy 481 CGAGTCGTGCGGCCCGGAGGCGCGTGTGGCGGACGAGCGCGCGGAGTGAAGGCCATG 540
Db 481 CGAGTCGTGCGGCCCGGAGGCGCGTGTGGCGGACGAGCGCGCGGAGTGAAGGCCATG 540
Qy 541 GACACCGCGCGCTCGCGCGTGTGCGGCGTACGTCGTCATCGTTCGCGGAGCGGG 600
Db 541 GACACCGCGCGCTCGCGCGTGTGCGGCGTACGTCGTCATCGTTCGCGGAGCGGG 600
Qy 601 ATGAGCTGTGCGCTCGCGCGCGGTGTACGAGGAAAGTGGAGTTCGTACAGCTG 660
Db 601 ATGAGCTGTGCGCTCGCGCGCGGTGTACGAGGAAAGTGGAGTTCGTACAGCTG 660
Qy 661 CGGAGCTCAAGGCGTGTTCGAGGACGAGACTCTGTTCCACTTCTACTCGGCGTCGATT 720
Db 661 CGGAGCTCAAGGCGTGTTCGAGGACGAGACTCTGTTCCACTTCTACTCGGCGTCGATT 720
Qy 721 CTTCTGGGCTATGATGCTGTCAGTCCAGACGCGGAGATGGGGTGGGGTGAACGGTG 780
Db 721 CTTCTGGGCTATGATGCTGTCAGTCCAGACGCGGAGATGGGGTGGGGTGAACGGTG 780
Qy 781 AAGCTGGTGGACTTTGCCCATGTGCCGAGGATGGGGTGATGGACCAACTTCTCTG 840
Db 781 AAGCTGGTGGACTTTGCCCATGTGCCGAGGATGGGGTGATGGACCAACTTCTCTG 840
Qy 841 GCGGGCTCTCTCGCTGATCAAGTTCGTTCTGACATTTGTCGAGACTCTCTCATACG 900
Db 841 GCGGGCTCTCTCGCTGATCAAGTTCGTTCTGACATTTGTCGAGACTCTCTCATACG 900
Qy 901 CAGCCTTTGGTCTCTTAAGAGAGGATCTCGCATTTGATTTGATAACAAAGCCCTA 960
Db 901 CAGCCTTTGGTCTCTTAAGAGAGGATCTCGCATTTGATTTGATAACAAAGCCCTA 960
Qy 961 CAAGTTTGTCTGAAAAGAGGCGCTCCGAGTTGTGCTGGGTGTGAGATCTGAGACG 1020
Db 961 CAAGTTTGTCTGAAAAGAGGCGCTCCGAGTTGTGCTGGGTGTGAGATCTGAGACG 1020
Qy 1021 GTGCTGCGCCACCTTGGTGTGCTTTCCTTTCCTGCAAAATACAGCAACCTGCT 1080
Db 1021 GTGCTGCGCCACCTTGGTGTGCTTTCCTTTCCTGCAAAATACAGCAACCTGCT 1080
Qy 1081 CTTTTCGCAACCCCTTACTTCGAGAAACTTTTTCCTTTCCTTTCCTTTCCTTTCG 1140
Db 1081 CTTTTCGCAACCCCTTACTTCGAGAAACTTTTTCCTTTCCTTTCCTTTCCTTTCG 1140
```

```
Qy 1141 TTACGTTGGATCTGTTTGTGCGACTCGGTGACAGGTTGTAAGCATGAGGAGCGCTGT 1200
Db 1141 TTACGTTGGATCTGTTTGTGCGACTCGGTGACAGGTTGTAAGCATGAGGAGCGCTGT 1200
Qy 1201 TGATCCGGCAACTGTGTGTCAGTCTTTGCGTGTGCTGCGTTTTCGATGCTTTTGCCTGC 1260
Db 1201 TGATCCGGCAACTGTGTGTCAGTCTTTGCGTGTGCTGCGTTTTCGATGCTTTTGCCTGC 1260
Qy 1261 TCGATCCGATGTGTACTGAGATCTGAGATGAGGAGTCTTACCTCCAAACGAATCCG 1320
Db 1261 TCGATCCGATGTGTACTGAGATCTGAGATGAGGAGTCTTACCTCCAAACGAATCCG 1320
Qy 1321 TCCGATAAAAAA 1344
Db 1321 TCCGATAAAAAA 1344

RESULT 2
AX513568
LOCUS 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 5 from Patent WO02059324.
ACCESSION AX513568
VERSION AX513568.1 GI:23559668
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 5 01-AUG-2002; (US)
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..922
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD52953.1"
/db_xref="GI:23559669"
/translation="MEDLPPEHQVAGHRASAKPGPLIDGSLFYKPLQAGDGEHE
VAFYEAFSAHAAPARIDTFFPRFHGRLLEPTEAQGPPEPHLVLDLLLAGREAPCV
ADIKGATWPPSPPEFYAKLMDRGTTSVLLGFRVSGRVVGPGEVWRTREPV
KAMDTAGVRRVLRVYSSVADEGMDCALAAAVYGGKGVLSQLRELKAMFEBQTLFPH
YSASILLGYDAAAFAAGGGGGVTKLVDPFAHVAEGDVIDHNFGLGSLIKFVSDI
VPETPQTQPLGPS"
ORIGIN
Query Match 64.4%; Score 865.4; DB 6; Length 923;
Best Local Similarity 98.8%; Pred. No. 6.1e-126;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 42 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACAAGTCGCGGCGCCGCGCTC 101
Db 43 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACAAGTCGCGGCGCTC 102
Qy 102 CGCAGCAAGCTGGGCGCGCTCATCGACGCTCGGCGCTCTTACAAAGCGCTCCAGGC 161
Db 103 CGCAGCAAGCTGGGCGCGCTCATCGACGCTCGGCGCTCTTACAAAGCGCTCCAGGC 162
Qy 162 CGGCGACCTGGGGAGCACAAGGTGCGCTTCTATAGGCGGTTCTCCGCCACGCGCGCT 221
Db 163 CGGCGACCTGGGGAGCACAAGGTGCGCTTCTATAGGCGGTTCTCCGCCACGCGCGCT 222
Qy 222 CGGCGCGCGCATCGGAGACCTTCTCCCGCGGTTCCAGGACGAGCTCTCTCCCCAC 281
Db 223 CGGCGCGCGCATCGGAGACCTTCTCTCCCGCGGTTCCAGGACGAGCTCTCTCCCCAC 282
Qy 282 CGAGGCGCAGCGCGGAGCGCGCATCTCACCTCGTCGACGACCTCTCTCGCGGGGTT 341
```

```
Db 283 CGAGCGCAGCCCGGGAGCGCATCCGACCTCGTCTCCGACGACCTCTCGGGGATT 342
Qy 342 TCAGCGCCCTGCGTCGAGACATCAAGATCGGCGCCATCACTGCGCCACCGATTGCGC 401
Db 343 TGAGCGCCCTGCGTCGAGACATCAAGATCGGCGCCATCACTGCGCCACCGATTGCGC 402
Qy 402 GGAGCCCTTACATCGCAAGTGCCTCGCCCAAGGACCGCGGGACCAAGCGTTCTGCTCGG 461
Db 403 GGAGCCCTTACATCGCAAGTGCCTCGCCCAAGGACCGCGGGACCAAGCGTTCTGCTCGG 462
Qy 462 ATTCCGCTCTCCGCGCTCGAGTGCCTCGGCGCCGAGGGCGCGTGTGGCGGACGAGCG 521
Db 463 ATTCCGCTCTCCGCGCTCGAGTGCCTCGGCGCCGAGGGCGCGTGTGGCGGACGAGCG 522
Qy 522 CCCGAGGTGAAGGCATGGAACCGCGCGGTCCGCGCGTGTCCGCGCGTACGTGTC 581
Db 523 CCCGAGGTGAAGGCATGGAACCGCGCGGTCCGCGCGTGTCCGCGCGTACGTGTC 582
Qy 582 ATCCGTTCCGACAGAGGGATGGAAGTGCCTCGCCCGCGGTGTACGAGGAAAGG 641
Db 583 ATCCGTTCCGACAGAGGGATGGAAGTGCCTCGCCCGCGGTGTACGAGGAAAGG 642
Qy 642 TGGAGTCTTGTACAGCTGCGGAGCTCAAGCGCGTGTTCGAGGAGCAGACTCTGTTCCA 701
Db 643 TGGAGTCTTGTACAGCTGCGGAGCTCAAGCGCGTGTTCGAGGAGCAGACTCTGTTCCA 702
Qy 702 CTTCTACTCGCGGTGATCTTCTGGGTATGATGCTGCTGCGAGCGGAGGAGA 761
Db 703 CTTCTACTCGCGGTGATCTTCTGGGTATGATGCTGCTGCGAGCGGAGG 762
Qy 762 TGGGGTGGGTGACGCTGAGCTGCTGAGCTTTCGCCATGTGGCGGAGGTGATGGGT 821
Db 763 TGGGGTGGGTGACGCTGAGCTGCTGAGCTTTCGCCATGTGGCGGAGGTGATGGGT 822
Qy 822 GATTGACCACAACTTCTCGCGGGCTCTGCTGCTGATCAAGTTCTGATTTG 881
Db 823 GATTGACCACAACTTCTCGCGGGCTCTGCTGCTGATCAAGTTCTGATTTG 882
Qy 882 TCCGAGACTCTCTATAGCGAGCTTTGGGTCTCTTTAAAG 922
Db 883 TCCGAGACTCTCTAGAGCGAGCTTTGGGTCTCTTTAAAG 923
```

RESULT 3  
AX513564  
LOCUS 1169 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 1 from Patent WO2059324.  
ACCESSION AX513564  
VERSION AX513564.1 GI:23559664

KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Shi, J., Beach, L. R., Wang, H., Rafalski, J. A. and Cahoon, R. B.  
TITLE Novel inositol polyphosphate kinase genes and uses thereof  
JOURNAL Patent: WO 02059324-A 1 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES  
source  
1..1169  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
84..806  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52951.1"  
/db\_xref="GI:23559665"  
/translation="MPDLHPPEHQVAGHRASKLGLPLIDGSLFYKPLQAGDRGEHE  
VAFYFAFSAHAAPVARIIRDTPFRPHGLRLPTEAPGPBPHPLVLDLLAGFEAPCV

ADIKIGAITWPPSPBPIAKYLAKDRGTTSLVLLGFRVLRPSRRPRRRVADGAPGGE  
GYRRRRPPPALRVLRITRGTARSRRCCTBEKVESCHSCASSRHGMRSLCST5  
TRRRFFWAMMLLOQQAEVGVG"

## ORIGIN

Query Match 62.9%; Score 844.8; DB 6; Length 1169;  
Best Local Similarity 97.2%; Pred. No. 1e-122;  
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

Qy 42 TCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCTC 101  
Db 74 TCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCTC 133  
Qy 102 CGCCAGCAAGTGGGCGCGCTCATCGACGGCTCCGGGCTCTTCTAAGCGCTCCAGGC 161  
Db 134 CGCCAGCAAGTGGGCGCGCTCATCGACGGCTCCGGGCTCTTCTAAGCGCTCCAGGC 193  
Qy 162 CGCGACGTCGGGAGCACGAGGTCGCTTCTATGAGGCGTCTCCGCCACCGCGCGT 221  
Db 194 CGCGACGTCGGGAGCACGAGGTCGCTTCTATGAGGCGTCTCCGCCACCGCGCGT 253  
Qy 222 CCGCGCGCATCGAGACACTTCTCCCGCGGTTCACGCGACGCGACTCTCTCCAC 281  
Db 254 CCGCGCGCATCGAGACACTTCTCCCGCGGTTCACGCGACGCGACTCTCTCCAC 313  
Qy 282 CGAGCGCAGCCCGGGAGCGCATCTCTCTCTCTCGACGACTCTCTCGCGGGTT 341  
Db 314 CGAGCGCAGCCCGGGAGCGCATCTCTCTCTCTCGACGACTCTCTCGCGGGTT 373  
Qy 342 TCAGCGCGCTGCTGCGAGACATCAAGATCGGCGCCATCAAGTGGCCACCGAGTTCGCC 401  
Db 374 TGAGCGCGCTGCTGCGAGACATCAAGATCGGCGCCATCAAGTGGCCACCGAGTTCGCC 433  
Qy 402 GGAGCGCTACATCCCAAGTACCTCGCCCAAGACCGCGGAGCGCTCTCTCG 461  
Db 434 GGAGCGCTACATCCCAAGTACCTCGCCCAAGACCGCGGAGCGCTCTCTCG 493  
Qy 462 ATTCCGCTCTCCGCGCTCGAGTCTCGCGCCCGAGGCGCGTGTGGCGAGCGAGCG 521  
Db 494 ATTCCGCTCT--TGCGTCCGAGTCTGCGCGCCCGAGGCGCGTGTGGCGAGCGAGCG 551  
Qy 522 CCGAGAGTGAAGCGCATGGAACCGCGCGGTCTCGCGCGCTCTCGCGCGCTACGTGTC 581  
Db 552 CCGAGAGTGAAGCGCATGGAACCGCGCGGTCTCGCGCGCTCTCGCGCGCTACGTGTC 611  
Qy 582 ATCCG--TTGCCGAGCGGGATGAGTCTGCGCTCGCGCGCGCTGTACGAGGAAAG 640  
Db 612 ATCCGCTTGGCGAGCGAGGGATGAGTCTGCGCGCTCGCGCGCGGTGTACGAGGAAAG 671  
Qy 641 GTGAGTCTTGTACAGCTGCGCGAGCTCAAGGGGTGTTGAGGAGCAGACTCTGTTCC 700  
Db 672 GTGAGTCTTGTACAGCTGCGCGAGCTCAAGGGATGTTGAGGAGCAGACTCTGTTCC 731  
Qy 701 ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGCTGAGTTCGAGCAGCGAG 760  
Db 732 ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGCTGAGTTCGAGCAGCGAG 791  
Qy 761 ATGGGGTGGGTGACCGTGAAGTGTGGGACTTTGCCCATGTGGCGCGGTGATGGG 820  
Db 792 GTGGGGTGGGTGAACAGTGAAGCTGTGGGACTTTGCCCATGTGGCGCGGTGATGGG 851  
Qy 821 TGATTGACCACAACTTCTCGCGCGGTCTGCTGCTGATCAAGTTCGTTCTGACATTG 880  
Db 852 TGATTGACCACAACTTCTCGCGCGGTCTGCTGCTGATCAAGTTCGTTCTGACATTG 911  
Qy 881 TTCCGGAGACTCTCTATACGAGCGCTTTGGGTCTCTTAAAGAGAGGATCTCGGCA--TTT 939  
Db 912 TTCCAGAGACTCTTTAGACGCGCTTTGGGTCTCTTAAAGAGAGGATCTCTGACATTTT 971  
Qy 940 CGATTGATACAAAG 955  
Db 972 TGATTGATACAAAG 987

```
RESULT 4
AX513566
LOCUS AX513566 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 3 from Patent WO02059324.
ACCESSION AX513566
VERSION AX513566.1 GI:23559666
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 3 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..736
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52952.1"
/db_xref="GI:23559667"
/translation="MPDLHPPEHQVAGHRASAKPGPLIDSGILFYKPLQAGDRGEHE
VAFYEAFSAHAAPVIRDTFFPRFHGTRLLPTEAQGPPEPHLVLDLLAGFEAPCV
ADIKIGALTWPPSPPEVIAKCLAMDRTTSLVLLGRVSGVRVVPGEAVRTPEPV
KAMDTGVRRVRLRYVSSACRBDGLRARGGVRRKXSLVTAARAGVVRGAASVPL
LLGVDSGL"
CDS
Query Match 62.3%; Score 837.4; DB 6; Length 923;
Best Local Similarity 97.5%; Pred. No. 1.5e-121;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
Qy 42 TCCCATACCATGTCGAGACACCTTTCTTCCCGGTTCCACGGACGGGACTCTCTCCAC 101
Db 43 TCCCATACCATGTCGAGACACCTTTCTTCCCGGTTCCACGGACGGGACTCTCTCCAC 102
Qy 102 CGCAGCAGCTGGGCGGCTCATCGAGCGCTCGGCTCTTCTACAGCGCTCCAGGC 161
Db 103 CGCAGCAGCTGGGCGGCTCATCGAGCGCTCGGCTCTTCTACAGCGCTCCAGGC 162
Qy 162 CGGCGACCGTGGGAGACAGAGTTCGCTTCTATGAGCGGTTCTCCGCGCACGCGCGT 221
Db 163 CGGCGACCGTGGGAGACAGAGTTCGCTTCTATGAGCGGTTCTCCGCGCACGCGCGT 222
Qy 222 CCGCGCGCATCGGACACCTTTCTTCCCGGTTCCACGGACGGGACTCTCTCCAC 281
Db 223 CCGCGCGCATCGGACACCTTTCTTCCCGGTTCCACGGACGGGACTCTCTCCAC 282
Qy 282 CGAGCGCAGCCCGGAGCGGATCTCTCACTCGTCTCTCGACACCTCTCGCGGGTT 341
Db 283 CGAGCGCAGCCCGGAGCGGATCTCTCGACACCTCTCTCGACACCTCTCTCGCGGAT 342
Qy 342 TCAGCGCGCTGCTCGCAGACATCAAGATCGGCGCATACGTTGGCCACCGAGTTCGCC 401
Db 343 TGAGCGCGCTGCTCGCAGACATCAAGATCGGCGCATACGTTGGCCACCGAGTTCGCC 402
Qy 402 GGAGCGCTTACATCGCCAGTGCCTCGCAGAGCCGGGACCAACGAGCGTTCTGCTCGG 461
Db 403 GGAGCGCTTACATCGCCAGTGCCTCGCAGTGCACCGGGGACCAACGAGCTTCTGCTCGG 462
Qy 462 ATTGGGCTTCCGGCGTCCAGTTCGTCGGCCCGGAGCGCGCTGTGGCGGACGAGCG 521
Db 463 ATTGGGCTTCCGGCGTCCAGTTCGTCGGCCCGGAGCGCGCTGTGGCGGACGAGCG 522
Qy 522 CCGGAGGTGAAGGCATGACACCGCGCGGCTCCGCGCGTGTCCGCGCGCTACGTGTC 581
```

```
Db 523 CCGGAGGTGAAGGCTATGGACACCGTGGCGTCCGCGCGTGTCTCGCGCTACGTGTC 582
Qy 582 ATCCG-TTGCAGAGGGGATGAGACTGTGCGTTCGCGCGCGGTTAGCGGAAAG 640
Db 583 ATCCGCTTGCAGAGGGGATGAGACTGTGCGTTCGCGCGCGGTTAGCGGAAAG 642
Qy 641 GTGGAGTCTTGTCACTGCTGCGGAGCTCAAGCGTGTTCGAGGAGCAGACTCTGTTC 700
Db 643 GTGGAGTCTTGTCACTGCTGCGGAGCTCAAGCGTGTTCGAGGAGCAGACTCTGTTC 702
Qy 701 ACTTACTCTGGGTCGATTTCTTCTGGGTCATGCTGCTGCAGTCGACGAGCGGAG 760
Db 703 ACTTACTCTGGGTCGATTTCTTCTGGGTCATGCTGCTGCAGTCGACGAGCGGAG 762
Qy 761 ATGGGGTGGGTCAGCTGTAAGCTGTGAGCTTTGCCATGTGGCGGAGGTGATGGG 820
Db 763 GTGGGGTGGGTAACAGTGAAGCTGTGAGCTTTGCCATGTGGCGGAGGTGATGGG 822
Qy 821 TGATTGACCAACTTCTCTGGGCGGCTCTGCTGCCTGATCAAGTTCGTTCTGACATTG 880
Db 823 TGATTGACCAACTTCTCTGGGCGGCTCTGCTGAGTCAAGTTCGTTCTGACATTG 882
Qy 881 TTCGGAGACTCTCTATACGACGCTTTGGGTCCTTCTTAA 921
Db 883 TTCAGAGACTCTCTAGAGCGAGCTTTGGGTCCTTCTTAA 923
RESULT 5
AX513583
LOCUS AX513583 3416 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 20 from Patent WO02059324.
ACCESSION AX513583
VERSION AX513583.1 GI:23559684
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 20 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..3416
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
72..407
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52959.1"
/db_xref="GI:23559685"
/translation="MPDLHPPEHQVAGHRASAKGLPIDSGILFYKPLQAGDRGEHE
VAFYEAFSAHAAPVIRDTFFPRFHGTRLLPTEAQGPPEPHLVLDLLAGFEAPCV
ADIKIGAIT"
CDS
Query Match 55.1%; Score 740.4; DB 6; Length 3416;
Best Local Similarity 90.2%; Pred. No. 2e-106;
Matches 862; Conservative 0; Mismatches 31; Indels 53; Gaps 4;
Qy 2 CAGAGTTCAGTCCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGCACC 61
Db 28 CACCATCGCCACCGCTCACCCCTTGTCTCCCATAG-----TCCCATACCATGTCGCACC 81
Qy 62 TCACCGCGCGGAGCACCAAGTCGCGCCACCGCGCTCGCCAGCAAGCTGGGCGCG 121
Db 82 TCACCGCGCGGAGCACCAAGTCGCGCGTCCCGCGCTCCGCGAGCAAGCTGGGCGCAC 141
Qy 122 TCATCGAGCGCTCGGCGCTTCTTACAAGCGCTTCCAGCGCGCTCCAGCGCGGAGCACG 181
```



Db 142 TCATCGACGACTCTGGCCCTCTTCTTACAAGCCGCTCCAGGCGGAGACCGTGGGGAGACG 201  
Qy 182 AGGTGCGCTTCTATGAGCGGTCTCCGCCACGCGCGCTCCGCCCGCGCATCCGAGACA 241  
Db 202 AGGTGCGCTTCTATGAGCGGTCTCCGCCACGCGCGCTCCGCCCGCGCATCCGAGACA 261  
Qy 242 CTTCTTCCCGGTTTCCAGCGCAGCGACTCTCTCCACCGAGCGCAGCCCGGGAGC 301  
Db 262 CTTCTTCCCGGTTTCCAGCGCAGCGACTCTCTCCACCGAGCGCAGCCCGGGAGC 321  
Qy 302 CGCATCTCTACCTCTCTCGACGACTCTCTCGCGGGTTTCAGCGCGCTCTCGTCGAG 361  
Db 322 CGCATCTCTCTCTCTCGACGACTCTCTCGCGGGTTTCAGCGCGCTCTCGTCGAG 381  
Qy 362 ACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCCGCGGAGCCCTACATCGCAAGT 421  
Db 382 ACATCAAGATCGTGCCATCAGT----- 405  
Qy 422 GCCTCGCAAGGACCGCGGAGCACGAGCGTTCTGCTCGGATTCGCGGTCTCCGCGTCC 481  
Db 406 -----GACCACGAGGTTCTGCTCGGATTCGCGGTCTCCGCGTCC 446  
Qy 482 GAGTCGTCCGCGCCGAGCGCGCTGTGCGGAGCGGCGCGCGAGGTGAAGGCCATGG 541  
Db 447 GAGTCGTCCGCGCCGAGCGCGCTGTGCGGAGCGGCGCGCGAGGTGAAGGCTATGG 506  
Qy 542 ACACCGCGGCTCCGCGCGCTCTCCGCGCTACGTGTCTATCCG-TTGGCGACGAGGG 600  
Db 507 ACATTTGTCCGCGCTCCGCGCGCTCTCCGCGCTACGTGTCTATCCGCTTCCGCGAGGG 566  
Qy 601 ATGAGTCTGCTCGCTCCGCGCGCTGTACGAGGAGAAAGGTGGAGTCTTTGTACAGCTG 660  
Db 567 ATGAGTCTGCTCGCTCCGCGCGCTGTACGAGGAGAAAGGTGGAGTCTTTGTACAGCTG 626  
Qy 661 CGCGAGCTCAAGCGCTGTGTGAGGAGCAGACTCTGTCCACTTCTATCTGCGCGTCAATT 720  
Db 627 CGCGAGCTCAAGCGCTGTGTGAGGAGCAGACTCTGTTCCTACTCTCTGCGCGTCAATT 686  
Qy 721 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 687 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746  
Qy 781 AAGCTGTGGTGAATTTGCCCATGTGCGAGGCTGATGGGTGATTTGACACCAACTTCTCTG 840  
Db 747 AAGCTGTGGTGAATTTGCCCATGTGCGAGGCTGATGGGTGATTTGACACCAACTTCTCTG 806  
Qy 841 GCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 807 GCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866  
Qy 901 GAGCTTTGGGTCCTTTTAAGAGAGGATCCTGGCA-TTTCGATTGTGATAACAAAG 955  
Db 867 GAGCTTTGGGTCCTTTTAAGAGAGGATCCTGGCATTTTCGATTGTGATAACAAAG 922

RESULT 6  
AX513578  
LOCUS AX513578 899 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 15 from Patent WO02059324.  
ACCESSION AX513578  
VERSION AX513578.1 GI:23559679

KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1  
Shi,J., Beach,L.R., Wang,H., Rafaleki,J.A. and Cahoon,R.E.  
Novel inositol polyphosphate kinase genes and uses thereof  
Patent: WO 02059324-A 15 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
Location/Qualifiers

source 1. .899  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
89\_424  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52958.1"  
/db\_xref="GI:23559680"  
/translation="MPDLHPPEHQVAGHRASAKGLPLIDSGGLFYKPLQAGDRGEHE  
VAFYEAFSAHAAPVAPARIRDITFFPRFHGTRLLPTEAQPGEPHPYLVLDLLAGFEAPCV  
ADIKIGAIT"  
ORIGIN  
Query Match 51.0%; Score 686; DB 6; Length 899;  
Best Local Similarity 88.7%; Pred. No. 7.5e-98;  
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;  
Qy 2 CACGAGTCAAGTCCGCTCACCCTCGGCCCATAGTCCCTTCCCACATACCATGTCGGACC 61  
Db 45 CACCATGCCACCCGTCACCCCTTGCTCCCATAG-----TCCCACATACCATGCCCGACC 98  
Qy 62 TCCACCCCGCGGAGCACCAAGTCCGCGCCACCGCCCTCCGCGAGCAAGCTGGGCCCGC 121  
Db 99 TCCACCCCGCGGAGCACCAAGTCCGCGCCCTCACCAGCGCTCCGCGAGCAAGCTGGGCCCGC 158  
Qy 122 TCATCGAGCGCTCCGCGCTCTTTCTACAAGCGCTCCAGCGCGGCGACCGTGGGGAGCAG 181  
Db 159 TCATCGAGCGCTCTGGCTCTTTCTACAAGCGCTCCAGCGCGGCGACCGTGGGGAGCAG 218  
Qy 182 AGGTGCGCTTCTATGAGCGGTCTTCGCGCCACCGCGCGCTCCGCGCCCGCATCCGAGACA 241  
Db 219 AGGTGCGCTTCTATGAGCGGTCTTCGCGCCACCGCGCGCTCCGCGCCCGCATCCGAGACA 278  
Qy 242 CTTCTTCCCGCGGTTCACGCGCAGGACTCTCTCCACCGAGCGCGAGCCCGCGGAGC 301  
Db 279 CTTCTTCCCGCGGTTCACGCGCAGGACTCTCTCCACCGAGCGCGAGCCCGGGAGC 338  
Qy 302 CGCATCTCTCACTCTGCTCTCGACGACTCTCTCGCGGGGTTCGAGCGCGCTCTGCTGCGAG 361  
Db 339 CGCATCTCTCACTCTGCTCTCGACGACTCTCTCGCGGGGTTCGAGCGCGCTCTGCTGCGAG 398  
Qy 362 ACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTTCGCGGAGCGCTTACATCGCCAACT 421  
Db 399 ACATCAAGATCGTGGCCATCAGT----- 422  
Qy 422 GCCTCGCAAGGACCGCGGAGCACGAGCGTTCGCTCGGATTCGCGCTCTCCGCGTCC 481  
Db 423 -----GACCATGAGCGATCTGCTCGGATTCACGCTCTCCGCGTCC 463  
Qy 482 GAGTCGTCCGCGCCCGAGGGCGCGGTGTCGCGAGCGAGCGCGCGAGGTGAAGGCGCATGG 541  
Db 464 GAGTCGTCCGCGCCCGAGGGCGCGGTGTCGCGAGCGAGCGCGCTGAGGTGAAGGCTATGG 523  
Qy 542 ACACCGCGGCGTCCGCGCGCTGCTTCGCGCGCTACGTGTCTATCCG-TTGGCGAGCGGG 600  
Db 524 ACATTGTCGCGCTCCGCGCGTGTCTCCGCGCTGCTCATGTCTGCTGCGCGAGGG 583  
Qy 601 ATGAGTCTGTCGCTCCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTTGTACAGCTG 660  
Db 584 ATGAGTCTGTCGCGCTCCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTTGTACAGCTG 643  
Qy 661 CGCGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCTCGCGTCAATT 720  
Db 644 CGCGAGCTCAAGGCGTGTTCGAGGGGAGACTCTGTTCCACTTCTACTCTCGCGTCAATT 703  
Qy 721 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 704 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763  
Qy 781 AAGCTGTGGTGAATTTGCCCATGTGCGCGAGGCTGATGGGGTGAATTCACCAACTTCTCTG 840  
Db 764 AAGCTGTGGTGAATTTGCCCATGTGCGCGAGGCTGATGGGGTGAATTCACCAACTTCTCTG 823





Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

REFERENCE AUTHORS TITLE	Sasaki, T., Matsumoto, T. and Yamamoto, K. Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC clone: P0415B12	CDS	/genes="P0415B12.2" /notes="start and end point are not identified" join(6065..6067,6850..7404) /gene="P0415B12.2" /notes="predicted by FGENSEH etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD25374.1" /db_xref="GI:49388256" /translation="MNSRDOIPLVYPRHEIVARLAKEDGGHQEQOQTVFVKGDGH SKNTLPSPASPSAATAPASALATIVASELPEEQNHALLPTVKAQILSAPKILSS SDTSTVTTTGTAAENSEGKTAVAAATVCMKSGLEKPTNPRVTRHVHVGSLGRL TVSQEGVGEGVLEGRGSCQITVPCH" complement(8669..15038) /gene="P0415B12.3" complement(join(8669..9060,9161..9334,9438..9569, 9640..9738,9929..10000,10080..10136,10206..10379, 10460..10554,10727..10826,10921..11073,11145..112275, 12870..12980,13078..13240,13613..13707,13793..14053, 14178..14331,14775..15038)) /gene="P0415B12.3" /notes="supported by full-length cDNA(s): AK121782" complement(join(8929..9060,9161..9334,9438..9569, 9640..9738,9929..10000,10080..10136,10206..10379, 10460..10554,10727..10826,10921..11073,11145..112275, 12870..12980,13078..13240,13613..13707,13793..14053, 14178..14338)) /gene="P0415B12.3" /notes="contains EST(s): AU082722(R4038), AU032417(R4038) contains full-length cDNA(s): AK121782" /codon_start=1 /product="gamma-tubulin complex component 5-like" /protein_id="BAD25375.1" /db_xref="GI:49388257" /translation="MVLPLNAGLMEADSGSFIKQLQSVSGDLPHAAPVPELSTQE HELVSFVEYQFGDTVLLYWDKTPGCEKAGIYVSHLSQTSRLAVLKPLFAATCL KQVELFVGRVSCGHGPTLTSAPASSVDWMLRLKALKEEQFLSVEFTITLLGL TDSMSISCSEAHLVQVQGVAPDAFWNSGAOMASSEVAHVAVNHFKLKEVCLD GESEYHMLAVIAGTLLPYQLCLDSWLYGILDDPNEEMFFYANKVATIDQPAFWM SYMVRVGRPTDSSSLADNESIRKELINQETATAAALLKSSNQGCADILCPVFLK IARALISAGKSFQLVHQVQTHRIQREVIFHEFNIDQNGTYISQKFRPDTSSIRIQD KREDIIEESTQFGNACKMDFLTLSSEFLICSLGLENGHDVDDYLRKLCDNAPVN PTVHKSNNVQETEVEEGENSEKTLKLRDATSGRDYDGMETKLNKAVMRDPTFV PGHDVSVTEVESYFNLSVYENPGITACQEMLRNKNKNSDLNISKSFLPLNDEN IKSIFGDRDSSGTIPGDTLSTVFPRLDGTDFKGFQDDSEVIROEDDRRTEALY TFPILPCVNVNPLSELPLQKDSLALSKALFKQSMRLRDLPLQVGILOECLSKCI KRVQDHKQILSKJMGDWRMLDFVIRAIYLLGSGMLQQLFTVTFDKLDKGNPMD DFELNTLQESIRNSADKMLTAPDSLVSVLAKHDTNDEETTSIRKKGQAQFGIE ALDVNFYKVSFPLDLIVNTEALKYNNQVMAFLKVKRAKFIIDETRKMKWGGST MHNFKOHLIVQKLLHFVDAPHQVMDRVVHSAWTELCDGWASATTLDEVMVEHAYL SSIQRCFVADKLMALIASRVKTIILGLALDFHNIETGLTGSTAPAKVACMEVD IKEQDFCVVFLRLILSFKLVNGHFFHLADLVTRINTYNYMSDSGSAIPGSRPR" complement(16514..23105) /gene="P0415B12.4" complement(join(16514..17069,17267..17644,18400..18594, 18674..18751,18816..19001,19077..19307,19386..19739, 19837..20034,20118..20288,20405..20512,20611..20829, 20914..21120,21227..21445,21564..21638,21998..22166, 22901..23105)) /gene="P0415B12.4" /notes="supported by full-length cDNA(s): AK065061" complement(join(16881..17069,17267..17644,18400..18594, 18674..18751,18816..19001,19077..19307,19386..19739, 19837..20034,20118..20288,20405..20512,20611..20829, 20914..21120,21227..21445,21564..21638,21998..22150)) /gene="P0415B12.4" /notes="contains EST(s): AU055859(S20105), AU085756(C52347), C27605(C52347), AU078722 (S20105) contains full-length cDNA(s): AK065061, AK063159" /codon_start=1 /product="RNA binding protein Rp120" /protein_id="BAD25376.1"
gene			
mRNA			
CDS			
gene			
mRNA			
CDS			
misc_feature			
gene			
mRNA			

FEATURES  
source

1..157419  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="P0415B12"  
complement(4313..5807)  
/gene="P0415B12.1"  
complement(join(4313..4547,5568..5807))  
/gene="P0415B12.1"  
/note="supported by full-length cDNA(s): AK121467"  
complement(4313..5807)  
/gene="P0415B12.1"  
/note="contains full-length cDNA(s): AK121467  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS  
probably inactive due to too long 3'UTR in CDS"  
6065..7404  
/gene="P0415B12.2"  
join(<6065..6067,6850..7404)





KEYWORDS	Zea mays	Novel inositol polyphosphate kinase genes and uses thereof
SOURCE	Zea mays	Patent: WO 02059324-A 18 01-AUG-2002;
ORGANISM	Zea mays	PIONEER HI-BRED INTERNATIONAL, INC. (US)
REFERENCE	1	
AUTHORS	Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.	
TITLE	Novel inositol polyphosphate kinase genes and uses thereof	
JOURNAL	PIONEER HI-BRED INTERNATIONAL, INC. (US)	
FEATURES	Location/Qualifiers	
source	1. .643	
	/organism="Zea mays"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:4577"	
ORIGIN		
Query Match	35.0%;	Score 470.4; DB 6; Length 643;
Best Local Similarity	95.9%;	Pred. No. 4.4e-64;
Matches	509; Conservative	0; Mismatches 18; Indels 4; Gaps 3;
Qy	42	TCCCATACATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGGCACCGGCCTC 101
Db	107	TCCCATACATGCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTCCACCGGCCTC 166
Qy	102	CGCCAGCAAGCTGGGCGCCCTCATCGACGGCTCCGGCCCTCTTCTACAAGCCGCTCCAGGC 161
Db	167	CGCCAGCAAGCTGGGCGCCCTCATCGACGGCTCCGGCCCTCTTCTACAAGCCGCTCCAGGC 226
Qy	162	CGCGCAAGCTGGGAGCACGAGTGCCTTCTATGAGCGTCTTCGCGCCACCGCCCGT 221
Db	227	CGCGCAAGCTGGGAGCACGAGTGCCTTCTATGAGCGTCTTCGCGCCACCGCCNGCT 286
Qy	222	CCGCGCCGATCCGAGACACTTCTTCCCGCGTTCACGCGAGCGACTCTCCCCAC 281
Db	287	CCGCGCCGATCCGAGACACTTCTTCCCGCGTTCACGCGAGCGACTCTCCCCAC 346
Qy	282	CGAGGCGCAGCGCGGAGCGCATCTCAGCTCGTCCTCGACGACTCTCTCGCGGGT 341
Db	347	CGAGGCGCAGCGCGGAGCGCATCTCAGCTCGTCCTCGACGACTCTCTCGCGGGT 406
Qy	342	TCAGGCGCCCTCGTTCGAGACATCAAGATCGCGCCATCACTGGGCCACCGAGTTGCC 401
Db	407	TGAGGCGCCCTCGTTCGAGACATCAAGATCGCGCCATCACTGGGCCACCGAGTTGCC 466
Qy	402	GGAGCCCTACATCGCAAGTGTCTGCCAAGACCGCGGACCAAGAGGTTCTGCTCGG 461
Db	467	GGAGCCCTACATCGNCAAGTACCTNGCCAAAGACCGCGGACCAAGAGGTTCTGCTCGG 526
Qy	462	ATTCGCGTCTCGGCGTCCGAGTCTCGGCGCCCGGAGGCGCGTGTGGCGGAGCGG 521
Db	527	ATTCGCGTCT--TGCGTCCGAGTCTCGGCGCCCGGAGGCGCGTGTGGCGGAGCGG 584
Qy	522	-CCGCGAGGTGAAGGCATGGACA-CCGCGCGGCTCCGCGCGTGTCTCGG 570
Db	585	CCCCGGGTGAAGCTATGGACACCCGTCGNGNCCGCGGTGCTTCGG 635
RESULT 11		
AX513581		
LOCUS	AX513581	519 bp DNA linear PAT 05-OCT-2002
DEFINITION	Sequence 18 from Patent WO02059324.	
ACCESSION	AX513581	
VERSION	AX513581.1	GI:23559682
KEYWORDS		
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	1	
AUTHORS	Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.	

TITLE	Novel inositol polyphosphate kinase genes and uses thereof	
JOURNAL	Patent: WO 02059324-A 18 01-AUG-2002;	
FEATURES	Location/Qualifiers	
source	1. .519	
	/organism="Zea mays"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:4577"	
ORIGIN		
Query Match	22.2%;	Score 298.4; DB 6; Length 519;
Best Local Similarity	89.4%;	Pred. No. 3.8e-37;
Matches	328; Conservative	0; Mismatches 38; Indels 1; Gaps 1;
Qy	627	GTACGGAGGAAAAGGTGGAGTCTTTGTACAGCTGCGGAGCTCAAGGGTGGTTCCAGGA 686
Db	2	GTACGGANGAANAANGTGGAGTCTTTGTACAGCTGCGGAGCTCAANGCGTGGTTCCAGGG 61
Qy	687	GCAGACTCTGTTCCACTTCTACTCTGGCGTCGATTTCTTGGGCTATGATGCTGCTGCAGT 746
Db	62	GCAGACTCTGTTCCACTTCTACTCTGGCGTCGATTTCTTGGGCTATGATGCTGCTGCAGT 121
Qy	747	CGCAGCAGCGGAGATGCGGGTGGGCTGACGCTGAAGCTGGTGGACTTTGCCCATGTGGC 806
Db	122	CGCAGCAGCGGANGTGGGGTGGGTTAACAGTGAAGCTGGTGGACTTTGCCCATGTGGC 181
Qy	807	CGAGGCTGATGGGGTGATTGACCAACTTCTCTGGCGGGCTCTGCTCGCTGATCAAGTT 866
Db	182	CGAGGCTGATGGGGTGATTGACCAACTTCTCTGGCGGGCTCTGCTTANCTGATCAAGTT 241
Qy	867	CGTTTCTGACATTTGTTCCGGAGACTCTCTATACGAGGCTTTGGGTCCTTCTTAAGAGAG 926
Db	242	TGTTTCTGACATTTGTTCCAGAGACTCTCTCAGACGCGAGCTTTGGGTCCTTCTTAAGAAA 301
Qy	927	GATCCTGGCA-TTTTCGATTGATAACAAGCCCTACAAGTTTGTCTGAAAAAAGAGCG 985
Db	302	GATCCTGGCAATTTTCGATTGATAACAAGGAANCATTTTCAGCTGCCAAAAAANCA 361
Qy	986	CCTCCGA 992
Db	362	CCAGTGA 368
RESULT 12		
AX513582		
LOCUS	AX513582	353 bp DNA linear PAT 05-OCT-2002
DEFINITION	Sequence 19 from Patent WO02059324.	
ACCESSION	AX513582	
VERSION	AX513582.1	GI:23559683
KEYWORDS		
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	1	
AUTHORS	Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.	
TITLE	Novel inositol polyphosphate kinase genes and uses thereof	
JOURNAL	Patent: WO 02059324-A 19 01-AUG-2002;	
FEATURES	Location/Qualifiers	
source	1. .353	
	/organism="Zea mays"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:4577"	
ORIGIN		
Query Match	14.3%;	Score 191.6; DB 6; Length 353;
Best Local Similarity	90.2%;	Pred. No. 2.1e-20;
Matches	238; Conservative	0; Mismatches 21; Indels 5; Gaps 3;
Qy	667	CTCAGGCGGTGGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGGCGTCAATTTCTG 726



Db 1 CTCAAGCATGGTTGGAGGACAGACTCTGTTCCACTTCTACTCGGGTGATTTCTCTG 60  
Qy 727 GGCTATGATGCTGCTGAGTGCAGAGCGGAGATGGGGTGGGGTGCAGGTAAGCTG 786  
Db 61 GGCTATGATGCTGCTGAGTGCAGAGCGGAGATGGGGTGGGGTGAACAGTGAAGCTG 120  
Qy 787 GTGACTTTGGCCATGTGGCCGAGGGTGAATGGGGT--GATTGACCACAACTTCTCTGGGCG 844  
Db 121 GTGACTTTGGCCATGTGGCCGAGGGTGAATGGGGTGAATTTGACCACTTCTCTGGGCG 180  
Qy 845 GGCTGCTGCTGCTGATCAAGTTC--GTTTCTGACATTTGTCGGGAGACTCTCTATAG-C 901  
Db 181 AGCTCTGCTAGTCTGATCAAGTTCGGTTCTTTGACATTTGTTCCAGANACTCTTAGAGGCC 240  
Qy 902 AGCCTTTGGTCTCTTAAGAGA 925  
Db 241 AGCCTTTGGTCTCTTAAAAA 264

RESULT 13  
AX513574  
LOCUS AX513574 1195 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 11 from Patent WO02059324.  
ACCESSION AX513574  
VERSION AX513574.1 GI:23559674  
KEYWORDS Eucalyptus grandis  
SOURCE Eucalyptus grandis  
ORGANISM Eucalyptus grandis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Myrtales; Myrtaceae; Eucalyptus.

REFERENCE  
1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.  
Novel inositol polyphosphate kinase genes and uses thereof  
Patent: WO 02059324-A 11 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
Location/Qualifiers  
FEATURES  
source  
1..1195  
/organism="Eucalyptus grandis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:71139"  
116..1048  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52956.1"  
/db\_xref="GI:23559675"  
/translation="MLKVPDQVAGHRGDGKGLVLDSDGRFYKPLQSDHRGDEVA  
FYESFYNTPEIPGHKRPAPHTKTIEASDGSPHPLVLEDLVSRGNPLMDIK  
TGSRTWPEASEEVIQKLEDRNSTSVSLGFRISGLRVYONSEAGFWQPKVYVSF  
NADGVRSLRPFVSNLSLGNVDPDCLYASKYVCHRGGILAQLLQKWEFEVQTNH  
FYSCSLIILDRESALDGCAPKPKVLVDFAHMDGHVIDHNFUGLGLCSVIKFIIRDIA  
DEDNKAACEYNLGLKENGIFYKSPELDEAC"

ORIGIN  
Query Match 13.7%; Score 184.6; DB 6; Length 1195;  
Best Local Similarity 55.6%; Pred. No. 2.3e-19;  
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;  
Qy 56 CCGACTCCACCGCGGAGCACCAGTCCGCGGCACCGCGCTCCCGCAGCAAGCTGG 115  
Db 114 CCATGCTCAAGTCCCGATCATCAAGTCGCGGTACCGGGAGACGGGGAAAGCTGG 173  
Qy 116 GCGCGCTCATGACGGCTCCGGCTCTTTCAAGCGCTCAGCGCCGCGGACCGTGGGG 175  
Db 174 GGCACACTGGTGAATTCGGGCGCGCTCTATAAGCCTCTCAGAGCGATCATCGCGAG 233  
Qy 176 AGCAGAGTGCCTCTTATGAGGGCTTCTCGCCACGCGCGCTCCCGCCGCACTCC 235  
Db 234 ACACGGAAGTGGCTTTTACGAGTCAATCTATTTCAATACCGAGATCCAGGTCACTC 293  
Qy 236 GAGACACCTCTTCCCGCGTTTCCACGCGCAGCGACTCTCCCAACCGAGCGAGCCCG 295  
Db 294 G---CAAAATCTTCTCGGTTTCAAGGAACTAAGACTAT---TGAGGCGTGTGATGAT 347

296 GCGAGCGCATCTCTACCTCGTCTCGACACCTCTCTCGCGGGTTTTCAGCGCGCTCGC 355  
348 CGGGTCTCTCAACCTCACTCTGGAGGATCTCGTCTCGGGTTCGACGAACCACTCTC 407  
356 TCGCAGACATCAAGATCGCGCGCATCACTGTTGGCCACCGAGTTCCGCGGAGCCCTACATCG 415  
408 TCATGGACATCAAGATCGGATCGGATCGAACAATGTTATCGGAGGCTCTGAGGAGTACATCC 467  
416 CCAAGTGCCTCGCCNAGGACCGCGGACACGAGCGTTCTGCTCGGATTCGGGCTCTCCG 475  
468 AAAAGTGTCTTAGAGAAAGATCGAAATAGCAACAAGCTTTTATGGGTTTATGGAATTTCTG 527  
476 GCGTCCGAGTCTCGG---CCCGAGGGCGCGTGTGGCGGAGCGGAGCGCCCGGAGGTGA 532  
528 GGCTAAGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAGAGTTGTTT 587  
533 AGGCCATGGACACCGCGCGCTCGCGCGCTGCTCGCGCGCTAGTGTCTATCC----- 585  
588 ATAGCTTTAATGCGGACCGTGTGAGTCTGAGTCTGAGGAAAGTTGTTTCTTCCAACTTGT 647  
586 -----GTTGCGGAGGAGGATGGACTGTGCGCTCGCGCGCGGCTGTACGAGGAAAG 640  
648 CTCGGGTCCAAATGTGGATCCGGATTTGTTGATGCATCAAAAGTTTACTGTACCGGG 707  
641 GTGAGTCTTGTCAACAGCTCGCGGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTC 700  
708 GTGGAATTTGGCAATTTGCTTCACTGAGGAGGAGTGTGAGTTTCAAGCAATATC 767  
701 ACTTCTACTCGCGCTGATTTCTTCTGGGCTATGATGCTGTCAGTTCGAGAGCGGAG 760  
768 ACTTCTATTCTGTTCACTCATTATCTATGACAGGAGTCTGCTTTGACCGC--- 823  
761 ATGGGGTGGGGTACCGTGAAGCTGTGACTTTGCGACTTTGCGCATGTGCGCGAGGATCGGG 820  
824 --TGTGCACACCCGAAAGTTAAACTGTTGAGTCTGTCATGATGATGATGATGATGATG 881  
821 TGATGACCACTTCTCGCGGGCTCTGCTGCTGATCAAGTCTGTTCTTCTGACATTG 880  
882 TGATGATCAAACTTCTTGGTGGCTCTGTTCTGTAATCAAGTTTATAGTGACATTG 941  
881 TTCGGA 887  
942 CTGATGA 948

RESULT 14  
LUS310150 25054 bp DNA linear PLN 24-JAN-2002  
LOCUS LUS310150  
DEFINITION Linum usitatissimum variety Bombay Ngc-D, Ngc-A and Ngc-B genes.  
ACCESSION AJ310150  
VERSION AJ310150.1 GI:13509206  
KEYWORDS Ngc-A gene; Ngc-A protein; Ngc-B gene; Ngc-B protein; Ngc-D gene;  
Ngc-D protein.  
SOURCE Linum usitatissimum (flax)  
ORGANISM Linum usitatissimum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Malpighiales; Linaceae; Linum.  
REFERENCE  
1 Dodds, P.N., Lawrence, G.J. and Ellis, J.G.  
Contrasting modes of evolution acting on the complex N locus for  
rust resistance in flax  
Plant J. 27 (5), 439-453 (2001)  
JOURNAL MEDLINE  
21461288  
PUBMED 11576428  
REFERENCE  
2 (bases 1 to 25054)  
Dodds, P.N.  
Direct Submission  
Submitted (22-MAR-2001) Dodds P.N., Plant Industry, CSIRO, GPO Box  
1600, Canberra, ACT 2601, Australia  
JOURNAL Location/Qualifiers  
FEATURES  
1..25054  
source





GTLVSSGFELTNILPSLSIYTKLRTLEVRSSQLPDLTNLKNLRLDLTTGCRELIEIA  
 GLHTLESLELSMERCPSVRKLDLAGIIKLTIHIHICTQLTEIRGLGLESIQMLPFM  
 SGQSIKEIPLNLSGLKNLKYFISKECRQLKVGNGLEELELDFNTDRRLKUKYLLKLV  
 SRKGOIVTQSGALFINN"

## ORIGIN

Query Match 13.5% Score 181.6 DB 8 Length 25054

Best Local Similarity	55.4%	Pred. NO. 4.9e-19;	
Matches 474;	Conservative	0;	Mismatches 349;
			Indels 33;
			Gaps 5;

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

50 CCATGTCCGACCTCCACCCGCGGAGCACCAAGTCGCCGCCACCGCGCTCCGCCAGCA 109

— — — — —

9605 CCAAAGCCATGTTCAAAGTTCGGAACATCAGGTTCCTGGACATCAAGCCATCAACGGCC 9664

110 AGCTGGGCGCCGCTCTATCGACCGGCTCCGCGCTCTCTCTATCAAGCGCGCTCCAGCGCGCGGACCC 169

y  
 110 AGCTGGGCCCCGTCATCGACGGCTCCGGCTCTCTACAAAGCCCGCTCAGGCCGGCGACC 189

b 9665 TCCTCGGTCCTCTTGTAGACGATTCAGGGCGATTCTACAAGCCACTTCAGGGCGATGAAC 9724

170 GTGGGAGCACGAGGTCGCCTTCTATGAGGCGTTCTCCGCCACGCCGCGTCCCGGCC y

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

9725 G TGGGGCAACCGAGATGGCATCTCTACACATCTTTTCTGGCTCACAAACGATCCACAAC 9784

230 GCATCCGAGACACCTTCTCTCCCTCCGCTCCACGGGACGGGACTCTCTCCCA CCGAGGGCG 289

A vertical scale from 0 to 100 in increments of 10, with horizontal tick marks extending to the right.

b 9785 ACGTCCG--GAGGTTTTCCTCCGGCCTTCCATGGCACTCAAAGTATAGAGGCTTCTGATG 9841

290 AGCCGGGAGCGCATCCTCACCTCGTCTCGACGACCTCCTCGGGGTTTCAGGCGC 349

[illegible]

D 9842 GATCTGG---CCCTCCACCCGCACCTTCATCTTAGAAGATCTCAGTTCGACTCGGTTCGATC 9898

350 CCTGCGTCCGAGACATCAAGATCGGCGCCATCACGTGGCCACCGAGTTCGCCGGAGCCCT 409

\_\_\_\_\_

b  
9899 CATGTGTCATGGACATCAAGATTGGTCCAGGACATGTTACCCGAGGCTTCTCAGCCT 9958

[illegible]

410 ACAATCGCCAAAGTGGCTTCGCCAAGGACCGCGGACCAAGAGCGTCTCTGCTCGGATTCGCG 469

**b**

9959 ACATCGAGAAATGCCTGAAGAAGGATGTGGAATCAAGCAGTCCCTTTCTCGGGTTTAGGA 1001

---

470 TCTCCGGCTCCGAGTCGTCCGCCCCCGA-----GGCGCGCTGTGGCGGACGG 517  
Y

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

b 10019 TATCCGGATTGCAGGTTACGGTAACGACAAAGAATCATCTGAGGTGGTCTTGAAGCCCTG 1007

518 AGCGCCCGGAGGTGAAGGCCATGGACACCGCCGGCGTCCGCCGCTACG 577

b 10079 AGAGGAAGCTTCTCCAGAACTCTTACTGCTGATGAAGTTAGGCTGGTGTCTTAAACGGTTTG 1013

RESULT 15

AX513572

QY	592	GACGAGGGGATGGACTGTGCGCTCCCGCGGGTGTACGGAGGAAAGGTGGAGTCTTG	651
Db	552	GATCATCATCCGATTGCGCTTTCGCAACGGAGGTCTAC-----GGCGCGTTTTG	602
QY	652	TCACAGCTGGCGGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCG	711
Db	603	GAGCGCTTGCAGAGCTCAAGGACTGGTTCGAGGTTTCAGACGGTGTATCACTTCTATTCT	662
QY	712	GGTCGATTCTTCTGGGCTATGATGCTGTCAGTCCGACGAGCGGAGATGGGGTGGG	771
Db	663	TGTTCTGTCTTGTGTGTACGAGAGGATCTAGGGAAA-----GGGAAAGCTACCAAC	716
QY	772	GTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGGTGTATGGGGTGAATGACCCAC	831
Db	717	CCTCTGGTCAAACTCGTTGACTTTGACACACGTGGTGGACGGAAACGGTGTCTATTGATCAC	776
QY	832	AACTTCCTGGGGGGCTCTGCTCGCTGATCAAGTTTCGTTTCTGACAT	878
Db	777	AACTTCCTGGGTGGCCTTGTTCCTTCATCAAGTTTCCTCAAGGATAT	823

Search completed: June 17, 2005, 00:32:05  
Job time : 5925 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2005, 00:32:13 ; Search time 879 Seconds  
(without alignments)  
9491.401 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1344	100.0	1344	14	US-10-042-894A-7
2	1307	97.2	1426	20	Sequence 7, Appli
3	865.4	64.4	923	14	Sequence 74438, A
4	844.8	62.9	1169	14	US-10-042-894A-5
5	837.4	62.3	923	14	US-10-042-894A-1
6	740.4	55.1	3416	14	US-10-042-894A-3
7	687.4	51.1	776	20	US-10-042-894A-20
					Sequence 74436, A

8	686	51.0	899	14	US-10-042-894A-15
9	545.6	40.6	1108	19	US-10-042-894A-15
10	470.4	35.0	643	14	US-10-042-894A-17
11	308	22.9	696	19	US-10-042-894A-18
12	298.4	22.2	519	14	US-10-042-894A-18
13	220.8	16.4	295	20	US-10-042-894A-19
14	191.6	14.3	353	14	US-10-042-894A-19
15	184.6	13.7	1195	14	US-10-042-894A-11
16	179	13.3	1105	14	US-10-042-894A-9
17	179	13.3	1345	18	US-10-042-894A-13
18	149.8	11.1	1020	14	US-10-042-894A-13
19	104.6	7.8	464	9	US-09-770-444-340
20	71.4	5.3	135638	16	US-10-314-657-1
21	71.4	5.3	135638	21	US-10-473-193-1
22	68.6	5.1	1383	18	US-10-417-700A-10
23	68.6	5.1	1478	19	US-10-437-963-64771
24	68.6	5.1	2007	19	US-10-437-963-57529
25	68.6	5.1	88421	9	US-09-976-059-1
26	68.2	5.1	2541	15	US-10-156-761-2855
27	68.2	5.1	9025608	15	US-10-156-761-1
28	67.2	5.0	1095	19	US-10-437-963-86085
29	66.8	5.0	1347	18	US-10-424-599-118038
30	66.2	4.9	546	19	US-10-437-963-5687
31	66.2	4.9	3471	15	US-10-156-761-2726
32	65.8	4.9	1491	20	US-10-425-115-143130
33	65.4	4.9	484	10	US-09-770-961-829
34	64.8	4.8	825	15	US-10-156-761-4706
35	64.8	4.8	9025608	15	US-10-156-761-1
36	64.4	4.8	1860	19	US-10-437-963-2917
37	64.2	4.8	1296	18	US-10-425-114-33515
38	64.2	4.8	1633	20	US-10-739-930-3065
39	63.8	4.7	5452	15	US-10-017-161-1481
40	63.8	4.7	5452	17	US-10-292-798-1189
41	63.6	4.7	987	15	US-10-156-761-3940
42	63.6	4.7	1656	19	US-10-437-963-50246
43	63.4	4.7	1581	15	US-10-156-761-3431
44	63.2	4.7	1227	21	US-10-399-693-15
45	63.2	4.7	138203	21	US-10-819-386A-1

ALIGNMENTS

RESULT 1

US-10-042-894A-7  
; Sequence 7, Application US/10042894A  
; Publication No. US2003009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US2003009011A1e1 Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 7  
; LENGTH: 1344  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (52)....(921)  
US-10-042-894A-7

Query Match 100.0%; Score 1344; DB 14; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
Qy	1	GCACGAGGTCA	GTCCGTCAC	CCCTCGCGCCCAT	AGTCCCTTCCCATACCAT	ATGTC	CGAC	60			
Db	1	GCACGAGGTCA	GTCCGTCAC	CCCTCGCGCCCAT	AGTCCCTTCCCATACCAT	ATGTC	CGAC	60			
Qy	61	CTCCACCCGCGGAG	CACCAAGTC	CGCGGCA	CCCGGCTCCGCCAG	CAGCT	TGGGCCG	120			
Db	61	CTCCACCCGCGGAG	CACCAAGTC	CGCGGCA	CCCGGCTCCGCCAG	CAGCT	TGGGCCG	120			
Qy	121	CTCATCGACGGCT	CCGGCTCTT	CTACAA	CGCGCTCCAGCGCGG	CACCGT	TGGGAGCAC	180			
Db	121	CTCATCGACGGCT	CCGGCTCTT	CTACAA	CGCGCTCCAGCGCGG	CACCGT	TGGGAGCAC	180			
Qy	181	GAGGTGCGCTT	CTATGAGGG	GTTC	CCGCCAC	AGCGCGCTCCGG	CCCGCAT	TCGAGAC	240		
Db	181	GAGGTGCGCTT	CTATGAGGG	GTTC	CCGCCAC	AGCGCGCTCCGG	CCCGCAT	TCGAGAC	240		
Qy	241	ACCTTCTTCCCGG	TTCCACG	CGGAC	CTCCTCC	CCACAGCGCG	CAGCCGGG	GAG	300		
Db	241	ACCTTCTTCCCGG	TTCCACG	CGGAC	CTCCTCC	CCACAGCGCG	CAGCCGGG	GAG	300		
Qy	301	CGGCATCCT	CACCTGCT	TCGAC	GACCTCCT	CGGGGT	TTTCAG	CGCGCT	TCGCA	360	
Db	301	CGGCATCCT	CACCTGCT	TCGAC	GACCTCCT	CGGGGT	TTTCAG	CGCGCT	TCGCA	360	
Qy	361	GACATCAAGAT	CGGCGCCAT	CACTG	TGGCCAC	CGAGTTCCG	CGGAGCC	TACATCG	CCCAAG	420	
Db	361	GACATCAAGAT	CGGCGCCAT	CACTG	TGGCCAC	CGAGTTCCG	CGGAGCC	TACATCG	CCCAAG	420	
Qy	421	TGCCTCGCCAA	GACCGCGG	ACCA	CGAGCGTT	CTGCTCG	GATTC	CGCGCT	CGCGCTC	480	
Db	421	TGCCTCGCCAA	GACCGCGG	ACCA	CGAGCGTT	CTGCTCG	GATTC	CGCGCT	CGCGCTC	480	
Qy	481	CGAGTCTCG	CGCGCGG	CGCGT	TGGCG	AGCGG	CGCGG	CGCTG	CAAGCCATG	540	
Db	481	CGAGTCTCG	CGCGCGG	CGCGT	TGGCG	AGCGG	CGCGG	CGCTG	CAAGCCATG	540	
Qy	541	GACACCGCG	CGGTGCT	CGCGCGT	CTCCGCGCT	TACGTGT	CATCCG	TTCG	CGACGAGGG	600	
Db	541	GACACCGCG	CGGTGCT	CGCGCGT	CTCCGCGCT	TACGTGT	CATCCG	TTCG	CGACGAGGG	600	
Qy	601	ATGSACTGT	CGCTCG	CGCGCGG	GTACG	GAGGAAA	AGGTG	GGAGTCT	TGTTCACAGCTG	660	
Db	601	ATGSACTGT	CGCTCG	CGCGCGG	GTACG	GAGGAAA	AGGTG	GGAGTCT	TGTTCACAGCTG	660	
Qy	661	CGCGAGCTCA	AGGCGTGT	TCGAGG	ACAGACT	CTGTT	CCACTT	CTACT	CGCGGT	CGATT	720
Db	661	CGCGAGCTCA	AGGCGTGT	TCGAGG	ACAGACT	CTGTT	CCACTT	CTACT	CGCGGT	CGATT	720
Qy	721	CTTCTGGGCTA	TGCTGCT	GAGTCG	AGCGGAG	ATGGGGT	TGGGTG	ACGGTG	ACGGTG	780	
Db	721	CTTCTGGGCTA	TGCTGCT	GAGTCG	AGCGGAG	ATGGGGT	TGGGTG	ACGGTG	ACGGTG	780	
Qy	781	AAGCTGGT	GGAATTT	GCCCAT	TGTG	CGCGAGG	GTGAT	GGGGT	TGATGACCA	CAACTT	840
Db	781	AAGCTGGT	GGAATTT	GCCCAT	TGTG	CGCGAGG	GTGAT	GGGGT	TGATGACCA	CAACTT	840
Qy	841	GGCGGGCT	CTGCTG	CTGATCA	AGTTCTG	TGACAT	TGTT	CCGAGACT	CTCTCAT	PACG	900
Db	841	GGCGGGCT	CTGCTG	CTGATCA	AGTTCTG	TGACAT	TGTT	CCGAGACT	CTCTCAT	PACG	900
Qy	901	CAGCCTTT	GGTCTCT	TAA	GAGAGAT	CCTG	GCATTT	CGATTT	TGATAT	CAAGCCCTA	960
Db	901	CAGCCTTT	GGTCTCT	TAA	GAGAGAT	CCTG	GCATTT	CGATTT	TGATAT	CAAGCCCTA	960
Qy	961	CAAGTTT	TGTCT	GAAA	AAAGAG	CGCCCT	CCGAGT	TGCTG	GGGTG	TGGAGATCT	1020
Db	961	CAAGTTT	TGTCT	GAAA	AAAGAG	CGCCCT	CCGAGT	TGCTG	GGGTG	TGGAGATCT	1020
Qy	1021	GTCTGCGG	CCACTT	GTGTT	GCCTT	GTGCTT	GCCTT	GTGCTT	GTGCTT	GTGCTT	1080
Db	1021	GTCTGCGG	CCACTT	GTGTT	GCCTT	GTGCTT	GCCTT	GTGCTT	GTGCTT	GTGCTT	1080

QY	1081	CCTTTTTTCG	CAACCCCTT	ACTTCCG	AAAGAAA	CTTTTTTTTT	TCCACCT	TTGGGGG	TT	CGA	1140				
DB	1081	CCTTTTTTCG	CAACCCCTT	ACTTCCG	AAAGAAA	CTTTTTTTTT	TCCACCT	TTGGGGG	TT	CGA	1140				
QY	1141	TTACGTTTGG	ATCTGGTT	TGTG	CCCACT	CGGT	CAGAG	TTGT	TAA	G	CATG	GAGGAG	CGGT	1200	
DB	1141	TTACGTTTGG	ATCTGGTT	TGTG	CCCACT	CGGT	CAGAG	TTGT	TAA	G	CATG	GAGGAG	CGGT	1200	
QY	1201	TGATCCGG	CAACTGT	GT	TCAGT	CTTT	TTCG	CTG	CCG	TTT	CT	GCAT	TGG	CT	1260
DB	1201	TGATCCGG	CAACTGT	GT	TCAGT	CTTT	TTCG	CTG	CCG	TTT	CT	GCAT	TGG	CT	1260
QY	1261	TGGCAT	CCGAT	GT	CTACT	GGAGAT	CGTAG	TGAT	G	AGAC	GTCT	TAC	CT	CC	1320
DB	1261	TGGCAT	CCGAT	GT	CTACT	GGAGAT	CGTAG	TGAT	G	AGAC	GTCT	TAC	CT	CC	1320
QY	1321	TCCGAT	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	1344
DB	1321	TCCGAT	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	1344
RESULT 2															
US-10-425-115-74438															
; Sequence 74438, Application US/10425115															
; Publication No. US20040214272A1															
; GENERAL INFORMATION:															
; APPLICANT: La Rosa, Thomas J.															
; APPLICANT: Kovalic, David K.															
; APPLICANT: Zhou, Yihua															
; APPLICANT: Cao, Yongwei															
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With															
; FILE REFERENCE: 38-21(5322)B															
; CURRENT APPLICATION NUMBER: US/10/425,115															
; CURRENT FILING DATE: 2003-04-28															
; NUMBER OF SEQ ID NOS: 369326															
; SEQ ID NO 74438															
; LENGTH: 1426															
; TYPE: DNA															
; ORGANISM: Zea mays															
; FEATURE:															
; OTHER INFORMATION: Clone ID: MRT4577_167886C.1															
US-10-425-115-74438															
Query Match 97.2%; Score 1307; DB 20; Length 1426;															
Best Local Similarity 99.5%; Pred. No. 0;															
Matches 1321; Conservative 0; Mismatches 5; Indels 1; Gaps 1;															
QY	8	GTCA	GTCCGT	CACCC	CTCG	CGCC	CCATAG	TCC	CTT	CC	ATAC	CATGTCC	GCAC	CT	67
DB	76	GTCA	GTCCGT	CACCC	CTCG	CGCC	CCATAG	TCC	CTT	CC	ATAC	CATGTCC	GCAC	CT	135
QY	68	CGCG	GAGCAC	CAAG	TC	CGCG	CCAC	CGCG	CTCG	CGC	AG	CAAG	CTCG	CGCG	127
DB	136	CGCG	GAGCAC	CAAG	TC	CGCG	CCAC	CGCG	CTCG	CGC	AG	CAAG	CTCG	CGCG	195
QY	128	ACGG	TC	CGGCTCT	TTT	TACA	AGCG	CTT	CCAG	CGCG	CGCA	CGTGG	GAGCAG	CAG	187
DB	196	ACGG	TC	CGGCTCT	TTT	TACA	AGCG	CTT	CCAG	CGCG	CGCA	CGTGG	GAGCAG	CAG	255
QY	188	CCTT	CTAT	GAG	CGT	TTCT	CGCC	CA	CGCG	CGGT	TC	CGG	CGCC	CGCAT	247
DB	256	CCTT	CTAT	GAG	CGT	TTCT	CGCC	CA	CGCG	CGGT	TC	CGG	CGCC	CGCAT	315
QY	248	TCCC	CGGT	TC	CAC	GGCAG	CGACT	CTCT	CC	CCAC	CGAG	CGCG	CAGCC	CGGGAG	307
DB	316	TCCC	CGGT	TC	CAC	GGCAG	CGACT	CTCT	CC	CCAC	CGAG	CGCG	CAGCC	CGGGAG	375
QY	308	CTCA	CT	GT	CT	CGAC	AG	CT	CT	CT	CG	CGGG	TTT	CAG	367
DB	376	CTCA	CT	GT	CT	CGAC	AG	CT	CT	CT	CG	CGGG	TTT	CAG	435

```
QY 368 AGATCGGCGCATACGTGCCACACGAGTTGCGCGGAGCCCTACATCGCAAGTGCCTCG 427
Db |||
QY 436 AGATCGGCGCATACGTGCCACACGAGTTGCGCGGAGCCCTACATCGCAAGTGCCTCG 495
Db |||
QY 428 CCAAGGACCGCGGACCAAGAGGTTCTGTCGGATTTCGCGGTCTCCGGTCCGAGTGC 487
Db |||
QY 496 CCAAGGACCGCGGACCAAGAGGTTCTGTCGGATTTCGCGGTCTCCGGTCCGAGTGC 555
QY 488 TGGGCCCCGAGGGCGCGTGTGGCGACGAGCGCGCGGAGGTGAAGGCCATGGACACCG 547
Db |||
QY 556 TGGGCCCCGAGGGCGCGTGTGGCGACGAGCGCGCGGAGGTGAAGGCCATGGACACCG 615
QY 548 CCGGCGTCCGCGCGTGCCTCCGCGCTACGTTGTCATCCGTTGCGGACGAGGGATGGACT 607
Db |||
QY 616 CCGGCGTCCGCGCGTGCCTCCGCGCTACGTTGTCATCCGTTGCGGACGAGGGATGGACT 675
QY 608 GTGCGTCCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTTGTCAAGCTGCGCGAGC 667
Db |||
QY 676 GTGCGTCCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTTGTCAAGCTGCGCGAGC 735
QY 668 TCAAGCGGTGTCGAGGACGAGACTCTGTTCCACTTCTACTCGGCGTGGATCTTCTCG 727
Db |||
QY 736 TCAAGCGGTGTCGAGGACGAGACTCTGTTCCACTTCTACTCGGCGTGGATCTTCTCG 795
QY 728 GCTATGATGCTGTCAGTCGACGAGCGGAGATGGGGTGGGTGACGGTGAAGCTGG 787
Db |||
QY 796 GCTATGATGCTGTTGACGTC-CAGCAGCCGAGATGGGGTGGGTGACGGTGAAGCTGG 854
QY 788 TGGACTTTGCCATGTGGCGAGGGTGATGGGGTGATTGACCAAACTTCTGGGGGGGC 847
Db |||
QY 848 TCTGCTCGCTGATCAAGTTGCTTTCTGACATGTTTCCGAGACTCTCTATACGACGCTT 907
Db |||
QY 915 TCTGCTCGCTGATCAAGTTGCTTTCTGACATGTTTCCGAGACTCTCTATACGACGCTT 974
QY 908 TGGGTCCTTTCTTAAGAGAGATCCTCGCATTTTCGATTTGATACAAAGCCCTACAAAGTTT 967
Db |||
QY 975 TGGGTCCTTTCTTAAGAGAGATCCTCGCATTTTCGATTTGATACAAAGCCCTACAAAGTTT 1034
QY 968 TGTCTGGAAGAAAGACGCTCCGAGTGTGTGCTGGGTGTGGAGATCTGAGACGGTCTGTCG 1027
Db |||
QY 1035 TGTCTGGAAGAAAGACGCTCCGAGTGTGTGCTGGGTGTGGAGATCTGAGACGGTCTGTCG 1094
QY 1028 GCCACTTGGTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1087
Db |||
QY 1095 GCCACTTGGTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1154
QY 1088 TCGCAACCCCTTACTTCCGAGAAACTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1147
Db |||
QY 1155 TCGCAACCCCTTACTTCCGAGAAACTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1214
QY 1148 GGATCTGGTTTGGCACTCGGTCAAGAGTTGTAAGCATGGAGGAGGGGTGTGTGATCCG 1207
Db |||
QY 1215 GGATCTGGTTTGGCACTCGGTCAAGAGTTGTAAGCATGGAGGAGGGGTGTGTGATCCG 1274
QY 1208 GCNACTGTGTGATCTTGGGCTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1267
Db |||
QY 1275 GCAACTGTGTGATCTTGGGCTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1334
QY 1268 CGATGTGTACTGAGATCGTAGTGATGAGCTCTTACCTCCCAAGCAATCCGTCGATG 1327
Db |||
QY 1335 CGATGTGTACTGAGATCGTAGTGATGAGCTCTTACCTCCCAAGCAATCCGTCGATG 1394
QY 1328 AAAAAA 1334
Db |||
QY 1395 AATCTAA 1401
```

## RESULT 3

US-10-042-894A-5

; Sequence 5, Application US/10042894A

; Publication No. US20030009011A1

```
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(922)
; US-10-042-894A-5
```

```
Query Match 64.4%; Score 865.4; DB 14; Length 923;
Best Local Similarity 98.8%; Pred. No. 2.8e-236;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 42 TCCCATACCATGTGCGACCTCACCGCGGAGACCAAGTGCCTCGGCGCACCGCGCTC 101
Db |||
QY 43 TCCCATACCATGTGCGACCTCACCGCGGAGACCAAGTGCCTCGGCGCACCGCGCTC 102
Db |||
QY 102 CGCGAGCAGCTGGCGCGCTCATCGACGGTCCGCGCTTCTTACAAGCGCTCCAGGC 161
Db |||
QY 103 CGCGAGCAGCTGGCGCGCTCATCGACGGTCCGCGCTTCTTACAAGCGCTCCAGGC 162
QY 162 CGCGAGCAGCTGGCGCGCTCATCGACGGTCCGCGCTTCTTACAAGCGCTCCAGGC 221
Db |||
QY 163 CGCGAGCAGCTGGCGCGCTCATCGACGGTCCGCGCTTCTTACAAGCGCTCCAGGC 222
QY 222 CGCGCGCGCATCTCGAGACACCTTCTTCCCGGTTTCCAGGACCGGACTCTCTCCAC 281
Db |||
QY 282 CGAGCGCAGCCCGGGAGCGCATCTCCTCGTCTCGACGACTCTCTCGCGGGT 341
Db |||
QY 283 CGAGCGCAGCCCGGGAGCGCATCTCCTCGTCTCGACGACTCTCTCGCGGGT 342
QY 342 TCAGGCGCCTCGCTCGCAGACATCAAGATCGCGCCATCAAGTGGCCACCGAGTTGCC 401
Db |||
QY 343 TCAGGCGCCTCGCTCGCAGACATCAAGATCGCGCCATCAAGTGGCCACCGAGTTGCC 402
QY 402 GGAGCCTTACATCGCCCAAGTGCCTGCGCAAGGACCGCGGACCAAGAGGTTCTGTCCG 461
Db |||
QY 403 GGAGCCTTACATCGCCCAAGTGCCTGCGCAAGGACCGCGGACCAAGAGGTTCTGTCCG 462
QY 462 ATTTCGCGTCTCGCGCTCCGAGTCTGCGCGCGCGCGCGCGCTGCGCGAGCG 521
Db |||
QY 463 ATTTCGCGTCTCGCGCTCCGAGTCTGCGCGCGCGCGCGCGCTGCGCGAGCG 522
QY 522 CCGGAGGTGAAGGCCATGGACACCGCGCGCTGCGCGCGCTGCTCCGCGCTACGTGTC 581
Db |||
QY 523 CCGGAGGTGAAGGCCATGGACACCGCGCGCTGCGCGCGCTGCTCCGCGCTACGTGTC 582
QY 582 ATCCGTTGCCGACGAGGGATGGAATGTGCGCTGCGCGCGCGGTGTACGAGGAAAGG 641
Db |||
QY 583 ATCCGTTGCCGACGAGGGATGGAATGTGCGCTGCGCGCGCGGTGTACGAGGAAAGG 642
QY 642 TGGAGTCTTGTCAAGCTGCGGAGCTCAAGCGGTGTTTCGAGGACGAGACTCTGTTC 701
Db |||
QY 643 TGGAGTCTTGTCAAGCTGCGGAGCTCAAGCGGTGTTTCGAGGACGAGACTCTGTTC 702
QY 702 CTCTACTCGCGTCAATTTCTTGGGCTATGATGCTGCTGTCAGTCCGAGCGGAGA 761
```

```
Db 703 CTTCTACTCGGCGTCGATTCCTCTGGGCTATGATGCTGCTGAGTCGAGCGAGCGGAGG 762
QY 762 TGGGGGTGGGGTGACGGTGAAGCTGGTGACCTTTGCCCCATGTGGCGAGGGGTGATGGGGT 821
Db 763 TGGGGGTGGGGTGACGGTGAAGCTGGTGACCTTTGCCCCATGTGGCGAGGGGTGATGGGGT 822
QY 822 GATTGACCAACAACCTTCTGGGGGGGCTCTGCTGCTGATCAAGTTTCGTTCTGACATTGT 881
Db 823 GATTGACCAACAACCTTCTGGGGGGGCTCTGCTGCTGATCAAGTTTCGTTCTGACATTGT 882
QY 882 TCCGAGAGACTCCTCATACACGAGCCTTTGGGTCTCTTTAAAG 922
Db 883 TCCAGAGACTCCTCAGACGACGCTTTGGGTCTCTTTAAAG 923

RESULT 4
US-10-042-894A-1
; Sequence 1, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(806)
US-10-042-894A-1

Query Match 62.9%; Score 844.8; DB 14; Length 1169;
Best Local Similarity 97.2%; Pred. No. 2.2e-230;
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 42 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCCTC 101
Db 74 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCCTC 133
QY 102 CGCCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGCGCTTCTTACAAGCCGCTCCAGCG 161
Db 134 CGCCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGCGCTTCTTACAAGCCGCTCCAGCG 193
QY 162 CGCGACCGTGGGAGCACAGGTGCGCTTCTATAGGGGCTTCTCGGCCACCGCGCGGT 221
Db 194 CGCGACCGTGGGAGCACAGGTGCGCTTCTATAGGGGCTTCTCGGCCACCGCGCGGT 253
QY 222 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCACGCGACGCACTCTCCCCAC 281
Db 254 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCACGCGACGCACTCTCCCCAC 313
QY 282 CGAGGCGCAGCCCGGGAGCCGATCCTCCTCCTCGACGACCTCCTCGCGGGGT 341
Db 314 CGAGGCGCAGCCCGGGAGCCGATCCTCCTCCTCGACGACCTCCTCGCGGGGT 373
QY 342 TCAGGCGCCTCGTTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCGCC 401
Db 374 TGAGGCGCCTCGTTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCGCC 433
QY 402 GGAGCCCTACATCGCCAAAGTGCTCGCCAAAGACCGCGGACACACGAGGTTCGTGCG 461
```

```
Db 434 GGAGCCCTACATCGCAAGTACCTCCCAAGGACCGCGGACCAAGAGGTTCTGCTCGG 493
QY 462 ATTCCGCGTCTCCGCGCTCCGAGTCGTTCGGCCCCAGGCGCGCTGTGCGACGAGCG 521
Db 494 ATTCCGCGTCT--TGCGTCCGAGTCGTTCGGCCCCAGGCGCGCTGTGCGACGAGCG 551
QY 522 CCCGAGGTGAAGCCATGGACACCGCGCGGTCCGCGCGCTGCTCCGCGCTACGTGTC 581
Db 552 CCCGAGGTGAAGGCTATGGACACCGTCGGCGTCCGCGCGTCTCCGCGCTACGTGTC 611
QY 582 ATCCG--TTGCCGACGAGGGGATGGACTGTGCGCTCCGCGCGCTGTACGAGGAAAAG 640
Db 612 ATCCGCTTCCGACGAGGGGATGGACTGTGCGCGCTCCGCGCGCTGTACGAGGAAAAG 671
QY 641 GTGGAGTCTTGTCAAGCTGCGGAGCTCAAGGCGTGTTCGAGGAGAGACTCTGTTCC 700
Db 672 GTGGAGTCTTGTCAAGCTGCGGAGCTCAAGGCAATGGTTGGAGGAGCAGACTCTGTTCC 731
QY 701 ACTTCTACTCGGCGTCGATTCCTTGGGCTATGATGCTGCTGACGTCGCGAGCGGAG 760
Db 732 ACTTCTACTCGGCGTCGATTCCTTGGGCTATGATGCTGCTGACGTCGCGAGCGGAG 791
QY 761 ATGGGGGTGGGTGACGCTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGCTGATGGG 820
Db 792 GTGGGGGTGGGTGAACAGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGCTGATGGG 851
QY 821 TGATTGACCAACAATTCCTGCGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTG 880
Db 852 TGATTGACCAACAATTCCTGCGCGAGCTCTGCTAGCTGATCAAGTTCGTTCTGACATTG 911
QY 881 TTCGGAGACTCCTCATAGCAGCCTTTGGGTCTTCTTAAGAGAGGATCCTGGCA-TTT 939
Db 912 TTCAGAGACTCCTTAGACGAGCCTTTGGGTCTTCTTAAGAGAGGATCCTGACATTTT 971
QY 940 CGATTGTATAACAAG 955
Db 972 TGATTGTATAACAAG 987

RESULT 5
US-10-042-894A-3
; Sequence 3, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(736)
US-10-042-894A-3

Query Match 62.3%; Score 837.4; DB 14; Length 923;
Best Local Similarity 97.5%; Pred. No. 2.7e-228;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 42 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCCTC 101
```



```
Db 43 TCCCATACCATGCCGAGCTCCACCGCGGAGCAACCAAGTCGCGGTCAACGCGGCTC 102
Qy 102 CGCAGCAAGCTGGGCGCGCTCATCGACGCTCGGCGCTCTTACAAAGCGCTCCAGGC 161
Db 103 CGCAGCAAGCTGGGCGCGCTCATCGACGCTCGGCGCTCTTACAAAGCGCTCCAGGC 162
Qy 162 CGGCGACCGTGGGAGCACAGGTCGCTTCTATAGGCGTCTCCGCCACGCGCGCT 221
Db 163 CGGCGACCGTGGGAGCACAGGTCGCTTCTATAGGCGTCTCCGCCACGCGCGCT 222
Qy 222 CCGGCGCGCATCCGAGACACCTTCTTCCCGCGGTTCACGCGACGCGACTCTCCCCAC 281
Db 223 CCGGCGCGCATCCGAGACACCTTCTTCCCGCGGTTCACGCGACGCGACTCTCCCCAC 282
Qy 282 CGAGGCGCGCGGAGCGGATCTCTCATCTGCTCTCGAGACCTCTCCGGGGTT 341
Db 283 CGAGGCGCGCGGAGCGGATCTCGCACCTCTGCTCTCGAGACCTCTCCGGGATT 342
Qy 342 TCAGGCGCGCTGCTCGACAGATCAAGATCGGCGCATCACGTTGGCCACCGAGTTCCGC 401
Db 343 TGAGGCGCGCTGCTCGACAGATCAAGATCGGCGCATCACGTTGGCCACCGAGTTCCGC 402
Qy 402 GGAGCCCTACATCGCCAAAGTGCCTTCGCAAGAGACCGCGGACCAACGAGCGTCTCTCGG 461
Db 403 GGAGCCCTACATCGCCAAAGTGCCTTCGCAAGAGACCGCGGACCAACGAGCGTCTCTCGG 462
Qy 462 ATTCCGCGCTTCGCGCGCTCGAGTCTGTCGCGCGCGCGGCGCGTGTGGCGAGCG 521
Db 463 ATTCCGCGCTTCGCGCGCTCGAGTCTGTCGTCGCGCGCGCGCGTGTGGCGAGCG 522
Qy 522 CCGGAGGTGAAGCCATGACACCGCGCGCTTCGCGCGCGTGTCTCGCGCGTACGTTGC 581
Db 523 CCGGAGGTGAAGCCATGACACCGCGCGCTTCGCGCGCGTGTCTCGCGCGTACGTTGC 582
Qy 582 ATCCG-TTGGCGAGCGAGGTGACTGTGCGCTTCGCGCGCGGTGTACGGAGGAAAG 640
Db 583 ATCCG-TTGGCGAGCGAGGTGACTGTGCGCTTCGCGCGCGGTGTACGGAGGAAAG 642
Qy 641 GTGAGTCTTGTCAAGCTCGCGAGCTCAAGCGTGTTCGAGGAGCAGACTCTGTTC 700
Db 643 GTGAGTCTTGTCAAGCTCGCGAGCTCAAGCGTGTTCGAGGAGCAGCTCTGTTC 702
Qy 701 ACTTCTACTCGCGTCAATTTCTGGGCTATGATGCTGCTGCAAGTGCAGCGAGCGGAG 760
Db 703 ACTTCTACTCGCGTCAATTTCTGGGCTATGATGCTGCTGCAAGTGCAGCGAGCGGAG 762
Qy 761 ATGGGGTGGGTGACCGTGAAGCTGGTGGACTTTGCGCATGTGGCGGAGGTGATGGG 820
Db 763 GTGGGGTGGGTGAACAGTGAAGCTGGTGGACTTTGCGCATGTGGCGGAGGTGATGGG 822
Qy 821 TGATTGACCAACTTCTCGGCGGCTCTGCTCGCTGATCAAGTCTCTTCTGACATTG 880
Db 823 TGATTGACCAACTTCTCGGCGGCTCTGCTAGCTGATCAAGTCTCTTCTGACATTG 882
Qy 881 TTCCGAGACTCTCTACACGAGCTTTGGGCTCTTTAA 921
Db 883 TTCCGAGACTCTCTACACGAGCTTTGGGCTCTTTAA 923
```

RESULT 6  
US-10-042-894A-20  
; Sequence 20, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafaleki, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286

```
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 3416  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (72)...(407)  
US-10-042-894A-20  
  
Query Match 55.1%; Score 740.4; DB 14; Length 3416;  
Best Local Similarity 90.2%; Pred. No. 1.7e-301;  
Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4;  
  
Qy 2 CACGAGGTCAAGTCGCTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCCGACC 61  
Db 28 CACATCGCACCCGTCAACCCCTTGTCCCATAG-----TCCCATACCATGTCCGACC 81  
Qy 62 TCACACCGCGGAGCACCAAGTCGCGGCGCACCGCGCTCCGCGAGCAAGCTGGGCGCG 121  
Db 82 TCACACCGCGGAGCACCAAGTCGCGGCTCACCGCGCTCCGCGCAGCAAGCTGGGCGCAC 141  
Qy 122 TCATCAGCGCTCCGCGCTCTTCTAAGCGCTCCAGCGCGGACCGTGGGAGCAGC 181  
Db 142 TCATCAGCGCTTCGCGCTCTTCTAAGCGCTCCAGCGCGGACCGTGGGAGCAGC 201  
Qy 182 AGTGTCCCTTCTATAGGCGTCTCCGCGCACCGCGCTCCGCGCGCGCATCCGAGACA 241  
Db 202 AGTGTCCCTTCTATAGGCGTCTCCGCGCACCGCGCTCTCCCGCGCGCATCCGAGACA 261  
Qy 242 CTTTCTTCCCGGTTCCAGCGCACCGACTCTCTCCCGACCGAGGCGCAGCCCGGGGAGC 301  
Db 262 CTTTCTTCCCGGTTCCAGCGCACCGACTCTCTCCCGCGCGCATCCGAGAGC 321  
Qy 302 CGATCTCTACCTGCTCTCGAGCAGCTCTCTCGGGGGTTTTCAGGCGCTTCTCGCTCGCAG 361  
Db 322 CGATCTCGACCTCTGCTCTCGAGCAGCTCTCTCGGGGGTTTTCAGGCGCTTCTCGCTCGCAG 381  
Qy 362 ACATCAAGATCGCGCGCATCACGTCGCGCACCGAGTTTCGCGGAGCCCTACATCGCCAAGT 421  
Db 382 ACATCAAGATCGGTGCGCATCACGT----- 405  
Qy 422 GCCTCGCAAGGACCGCGGACCCAGCGACTTCTGCTCGGATTCGCGCTCTCGCGCGTCC 481  
Db 406 -----GACCAACGAGCGTCTGCTCGGATTCGCGCTCTCGCGCGTCC 446  
Qy 482 GAGTCTGCGCGCGCGGCGCGCTGTGGCGAGCGAGCGCCCGAGGCTGAAGCCATGG 541  
Db 447 GAGTCTGCGCGCGCGGCGCGCTGTGGCGAGCGAGCGCCCGAGGTTGAAGGCTATGG 506  
Qy 542 ACACCGCGCGCTCGCGCGCTCTCGGGCGTACGTGTCATCCG-TTGGCGAGCGAGGG 600  
Db 507 ACATTGTGCGCGCTCGCGCGCTCTCGGGCGTACGTGTCATTCGCTTGGCGAGCGAGGG 566  
Qy 601 ATGGACTGTGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTGTCAAGCTG 660  
Db 567 ATGGACTGTGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTGTCAAGCTG 626  
Qy 661 CGCGAGCTCAAGCGGTGGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGCGGTCCATT 720  
Db 627 CGCGAGCTCAAGCGGTGGTTCGAGGAGGCGAGACTCTGTTCACCTTCTACTCGCGGTCCATT 686  
Qy 721 CTTCTGGGTATGATGCTGTCAGTCGAGTCGAGAGGCGAGATGGGGTGGGCTGACCGTG 780  
Db 687 CTTCTGGGTATGATGCTGTCAGTCGAGTCGAGAGGCGAGGTTGGGGTGGGTTAAGCTG 746  
Qy 781 AAGCTGTGAGCTTTCGCCCATGTGGCGCGGAGGTGATGGGGTGTGATGACCAACTTCTCTG 840  
Db 781 AAGCTGTGAGCTTTCGCCCATGTGGCGCGGAGGTGATGGGGTGTGATGACCAACTTCTCTG 840
```

Db 747 AAGCTGGTGGACTTTGCCCATGTGGCGAGGGTGATGGGGTGATTGACCAACAACATTCCTG 806  
Qy 841 GGGGGGCTCTGCTCGCTGATCAAGTTTCGTTTCGACATTTGTTCCGGAGACTCTCTCATAG 900  
Db 807 GGGGGGCTCTGCTAGCTGATCAAGTTTGTTCGACATTTGTTCCAGAGACTCTCTCAGACG 866  
Qy 901 CAGCCCTTTGGGTCTCTTTAAAGAGAGAGTCCCTGGCA-TTTTCGATTTTGATAACAAG 955  
Db 867 CAGCCCTTTGGGTCTCTTTAAAGAGAGAGTCCCTGGCA-TTTTCGATTTTGATAACAAG 922

RESULT 7  
US-10-425-115-74436  
; Sequence 74436, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 74436  
; LENGTH: 776  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unaure  
; LOCATION: (1)..(776)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_167884C.1  
US-10-425-115-74436

Query Match 51.1%; Score 687.4; DB 20; Length 776;  
Best Local Similarity 99.7%; Pred. No. 1.6e-185;  
Matches 699; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 566 TCCGGGCTACGTTGTCATCGTTGCGGACGAGGGATGGACTGTGGCTCGCGGGCGG 625  
Db 77 TCCGGGCTACGTTGTCATCGTTGCGGACGAGGGATGGACTGTGGCTCGCGGGCGG 136  
Qy 626 TGACGGAGGAAAAGTGGAGTCTTGTACAGCTGCGGAGCTCAAGGCGTGGTTCGAG 685  
Db 137 TGTACGGAGGAAAAGTGGAGTCTTGTACAGCTGCGGAGCTCAAGGCGTGGTTCGAG 196  
Qy 686 AGCAGACTCTGTCCACTTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGCGAG 745  
Db 197 AGCAGACTCTGTTCCTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGTCGAG 256  
Qy 746 TGCAGCAGCGAGATGGGGTGGGTCGACGCTGAAGCTGGTGGACTTTGCCCATGTGG 805  
Db 257 TGCAGCAGCGAGATGGGGTGGGTCGACGCTGAAGCTGGTGGACTTTGCCCATGTGG 316  
Qy 806 CCGAGGGTGATGGGGTGATTTGACCAACAACCTCTCGGCGGGCTCTGCTCGCTGATCAAGT 865  
Db 317 CCGAGGGTGATGGGGTGATTTGACCAACAACCTCTCGGCGGGCTCTGCTCGCTGATCAAGT 376  
Qy 866 TCGTTTCTACATTTGTCGGAGACTCTTCATACGCGAGCTTTGGGCTCTTTTAAGAGA 925  
Db 377 TCGTTTCTACATTTGTCGGAGACTCTTCATACGCGAGCTTTGGGCTCTTTTAAGAGA 436  
Qy 926 GGATCCTGGCATTTTCGATTTGATAAAGCCCTACAAGTTTGTCTGTAAGAAAGAGCG 985  
Db 437 GGATCCTGGCATTTTCGATTTGATAAAGCCCTACAAGTTTGTCTGTAAGAAAGAGCG 496  
Qy 986 CTTCCGAGTTGTCTGGGTGTGGAGATCTGAGACGCTCGTCCGCCCACTTGGTTGCTTGG 1045  
Db 497 CTTCCGAGTTGTCTGGGTGTGGAGATCTGAGACGCTCGTCCGCCCACTTGGTTGCTTGG 556

## RESULT 8

US-10-042-894A-15  
; Sequence 15, Application US/10042894A  
; Publication No. US2003009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US2003009011A1el Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 899  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (89)....(424)  
US-10-042-894A-15

Query Match 51.0%; Score 686; DB 14; Length 899;  
Best Local Similarity 88.7%; Pred. No. 4e-185;  
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;  
Qy 2 CACGAGGTGAGTCCGTCAACCCCTCGCGGCATAGTCCCTTCCCATACCATGTCCGACC 61  
Db 45 CACCATCGCCACCCGTCAACCCCTTGTCTCCCATAG-----TCCCATATACCATGCCGACC 98  
Qy 62 TCACCCCGCGGAGACCAAGTCGCGGCACCGCGCTCCGCCAGCAAGCTGGGGCCGC 121  
Db 99 TCACCCCGCGGAGACCAAGTCGCGGCACCGCGCTCCGCCAGCAAGCTGGGGCCGC 158  
Qy 122 TCATCGAGGGTTCGGGCTCTTTCTAAGCCGCTCCAGCCGCGGACCGTGGGGAGCAG 181  
Db 159 TCATCGAGGGTTCGGGCTCTTTCTAAGCCGCTCCAGCCGCGGACCGTGGGGAGCAG 218  
Qy 182 AGGTGCGCTTCTATGAGGGTTCCTCGGCCAGCGCGCTCCCGGCCCGCATCCGAGACA 241  
Db 219 AGGTGCGCTTCTATGAGGGTTCCTCGGCCACCGCGCGCTCCCGGCCCGCATCCGAGACA 278  
Qy 242 CTTCTTCCCCCGGTTCCACGGCAGCGACTCTCTCCACACCGAGGCGAGCCCGGGGAGC 301  
Db 279 CTTCTTCCCCCGGTTCCACGGCAGCGACTCTCTCCACACCGAGGCGAGCCCGGGGAGC 338  
Qy 302 CGCATCTCACCTCGTCTCTCGACGACTCTCTCGCGGGGTTTCAGGCGCCCTCGCTGCGAG 361  
Db 339 CGCATCTCGTACCTCGTCTCTCGACGACTCTCTCGCGGGGTTTCAGGCGCCCTCGCTGCGAG 398

```
QY 362 ACATCAAGATCGCGCATCAGTGGCCACCGAGTTTCGCGGAGCCTTACATCGCCAGT 421
Db |||||
QY 399 ACATCAAGATCGGTGCATCAGCT----- 422
Db |||||
QY 422 GCCTGCCAAGGACCGCGGACACAGAGCGTTCTGCTCGGATTCGCGTCTCCGGCGTCC 481
Db |||||
QY 423 -----GACCATGAGCGATCTGCTCGGATTCACGCTCCGGCGTCC 463
Db |||||
QY 482 GAGTCGTGGCCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGGAGGTGAAGGCCATGG 541
Db |||||
QY 464 GAGTCGTGGCCCCGAGGGCGCGTGTGGCGGACGAGCGCCCTGAGGTGAAGGCTATGG 523
Db |||||
QY 542 ACACCCCGCGGTCCGCGCGTCTCCGGCGTACGTGTGCATCGG-TTGCAGACGAGGGG 600
Db |||||
QY 524 ACATTGCGGCGTCCGCGCGTCTCCGGCGTGCATGTCTCCGCGGAGGGG 583
Db |||||
QY 601 ATGGAAGTGTGCGTCCGCGCGGTGTACGAGGAAAGGTGGAGTCTTTGTCAACAGCTG 660
Db |||||
QY 584 ATGGAAGTGTGCGTCCGCGCGGTGTACGAGGAAAGGTGGAGTCTTTGTCAACAGCTG 643
Db |||||
QY 661 CGGAGCTCAAGCGGTGGTTCGAGGACAGACTCTGTTCACATTCTACCTCGGCGTCAAT 720
Db |||||
QY 644 CGGAGCTCAAGCGGTGGTTCGAGGGGACAGACTCTGTTCACATTCTACCTCGGCGTCAAT 703
Db |||||
QY 721 CTTCTGGGTATGATGCTGCTGACGTGCGAGCGGAGAGTGGGGTGGGGTGACGGTG 780
Db |||||
QY 704 CTTCTGGGTATGATGCTGCTGACGTGCGAGCGGAGGAGTGGGGTGGGGTAAACAGTG 763
Db |||||
QY 781 NAGCTGTGAGCTTGGCCATGTGGCGAGGGTGTGGGGTGTGACCAACATTCCTG 840
Db |||||
QY 764 NAGCTGTGAGCTTGGCCATGTGGCGAGGGTGTGGGGTGTGACCAACATTCCTG 823
Db |||||
QY 841 GCGGGCTGTGCTGCTGATCAAGTTGTTTCTGACATGTTCCGAGAGCTCTCTCATAG 900
Db |||||
QY 824 GCGGGCTGTGCTGATGATCAAGTTGTTTCTGACATGTTCCAGAGACTCTCTTAGACG 883
Db |||||
QY 901 CAGC 904
Db |||||
QY 884 CAGC 887
Db |||||
```

RESULT 9  
US-10-437-963-4650  
; Sequence 4650, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 4650  
; LENGTH: 1108  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_11514C.1  
US-10-437-963-4650

Query Match 40.6%; Score 545.6; DB 19; Length 1108;  
Best Local Similarity 78.6%; Pred. No. 4.7e-145;  
Matches 695; Conservative 0; Mismatches 174; Indels 15; Gaps 3;

```
QY 55 TCCGACCTCCACCCCGCGGAGCACCAAGTTCGCGGCCACCGCGCTCCCGACCAAGCTG 114
Db |||||
QY 7 TCCGACCTCCGCGCCCGCGGAGCACAGGTGGGGGACCGCGCTCCCGCAAGCTG 66
Db |||||
QY 115 GCGCGCTCATCGAGGCTCCGGCTCTTCTACAGCCGCTCCAGCCCGCGGACCGTGGG 174
Db |||||
QY 67 GCGCGCTCTGTCGAGCGGCGAGGGCTCTTCTACAAGCCCTCCAGCCCGGGAGCGGG 126
Db |||||
QY 175 GAGCAGAGTCCGCTTCTATGAGGCTTCTCCGCCACCGCGCTCCCGGCCGCGATC 234
Db |||||
QY 127 GAGCAGAGGCGCGCTTCTACGCGGCTTCAACCGCACCGGCGCTCCCGCCCGGGTC 186
Db |||||
QY 235 CGAGACACCTTCTCCCGCGGTTCACAGCGACGCGACTCTCTCCCAACCGAGGCGCAGCC- 293
Db |||||
QY 187 CGGGGCGCTTCTTCCCGGCTTCCACGCGACCGCGCTTCTTCCCGCCCGCAGCGCCC 246
Db |||||
QY 294 --CGGGGAGCGCGATCTCCTCCTGCTCGAGGACCTCTCTCGGGGGTTCAGGGGCCC 351
Db |||||
QY 247 GCGCGCGCGCTTACCGCGACATCTGCTCTCGAGGACCTCTCTCGGGGCTTCCGTTCCCC 306
Db |||||
QY 352 TCGGTGCGAGACATCAAGATCGCGGCTATCGTGGCCACCGAGTTTCGCGGAGCCCTAC 411
Db |||||
QY 307 TCGGTGCGGAGCTCAAGATCGCGCTGACGCTGGCGCGCGGATCCCGGAGCCCTAC 366
Db |||||
QY 412 ATCGCCTGCTCCGCAAGGACCGCGGACCAAGAGGCTTCTGCTCGGATTCGCGCTC 471
Db |||||
QY 367 GTCGCGAGTCTGCTCGCAAGGACCGCGAGACCAACGAGCGCTCTCTCGGCTTCGCGCTC 426
Db |||||
QY 472 TCGGGCTCCGAGTCTGTC--GGCCCGAGGGCGCGTGTGGCGGACGAGGCGCGCGAG 528
Db |||||
QY 427 TCGGGCTCCGAGTCTGATGCTCGGGGCGCGCTGTGGCGCGCGGACCGGTCGAG 486
Db |||||
QY 529 GTGAGGCCATGAGACCGCGCGCTCCGCGCTCTCGCGCTACGTTGTCATCGTT 588
Db |||||
QY 487 CTGAGGAGATCGAGCGCGCGGCTCCGCGCTCTCGCGCTACGTTGTCACGCGC 546
Db |||||
QY 589 GCGCAGAGGGGATGAGTCTGCTGCGCTCGCGCGCGGCTGTACGAGAGAAAGGTGGAGTC 648
Db |||||
QY 547 GCGCGGAGCGCTGAGCTCGCGCTCGCGCGCGGCTGTACGAGGGGAGCGCGCTC 606
Db |||||
QY 649 TTGTACAGCTGCGGAGCTCAAGGCTGGTTCGAGGAGGAGACTCTGTTCCACTTCTAC 708
Db |||||
QY 607 CTGGCTCAGCTCGGGAGCTCAAGGCTGGTTCGAGGAGCAAAACCTGTATCCACTTCTAC 666
Db |||||
QY 709 TCGGGCTCGATTTCTTCTGGCT-----ATGATGCTGCTGAGTCGCGAGGAGCGGA 759
Db |||||
QY 667 TCGGGCTCGATTTCTTCTGGCTACGAGCGCAATGCGGCGCGGCTGCTCCCGAGGT 726
Db |||||
QY 760 GATGGGGTGGGGTGACGGTGAAGCTGGTGGACTTTTGGCCATGTGGCGGAGGTGATGGG 819
Db |||||
QY 727 GGNAGCGGCGGTGAAGGTGAAGCTGGTGGACTTCGCGCATGTGAGGATGGGAGCGG 786
Db |||||
QY 820 GTGATTGACCAACTTCTTCCGCGGCTGCTGCTGCTGATCAAGTTCTGTTCTGACATT 879
Db |||||
QY 787 GTGATTGACCAACTTCTTGGCGGCTCTGCTGCTCATCAAGTTCTATCGGCGACATT 846
Db |||||
QY 880 GTTCCGAGACTCTCATACGAGCCTTTGGGCTCTTCTTAAGA 923
Db |||||
QY 847 GTCGCGGAGGTACCGAGAGGCGCTCTTCAGATCATTTCTTGAAG 890
Db |||||
```

RESULT 10  
US-10-042-894A-17  
; Sequence 17, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1  
; TITLE OF INVENTION: Genes and Uses Thereof

; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 643  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(643)  
; OTHER INFORMATION: n = A, T, C or G  
US-10-042-894A-17

Query Match 35.0%; Score 470.4; DB 14; Length 643;  
Best Local Similarity 95.9%; Pred. No. 1.1e-123;  
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;  
QY 42 TCCCATACATGTCGACCTCCACCCCGCGAGCACCAAGTCGCGCCACCCGCGCTC 101  
DB 107 TCCCATACATGTCGACCTCCACCCCGCGAGCACCAAGTCGCGCTC 166  
QY 102 CGCCAGCAAGTCGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 161  
DB 167 CGCCAGCAAGTCGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 226  
QY 162 CGCCAGCAAGTCGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 221  
DB 227 CGCCAGCAAGTCGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 286  
QY 222 CGCCAGCAAGTCGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 281  
DB 287 CGCCAGCAAGTCGGCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 346  
QY 282 CGAGCGGAGCCCGGGAGCGCATCTTCTTCCCGGTTTCCAGCGACGCGACTCTTCCCGGC 341  
DB 347 CGAGCGGAGCCCGGGAGCGCATCTTCTTCCCGGTTTCCAGCGACGCGACTCTTCCCGGC 406  
QY 342 TCAGGCGCCCTCGCTCGCAGACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCGCC 401  
DB 407 TGAGGCGCCCTCGCTCGCAGACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCGCC 466  
QY 402 GGAGCCCTACATCGCAAGTCGCTGCCAAGACCGCGGACCAAGAGCGTTTCTGCTCGG 461  
DB 467 GGAGCCCTACATCGCNAAGTACTTNGCCAAAGACCGCGGACCAAGAGCGTTTCTGCTCGG 526  
QY 462 ATTCCGCGTCTCGGCGTCCGAGTCGTCGCGCCCGAGCGCGCGTGTGCGGACCGAGCG 521  
DB 527 ATTCCGCGTCT--TGCGTCCGAGTCGTCGCGCCCGAGCGCGCGTGTGCGGACCGAGCG 584  
QY 522 -CCCGAGGTGAAGGCCATGGACA--CCGCGCGCTCCGCGCGTGTCTCCGG 570  
DB 585 CCCCCGGGTGAAGCTATGGACACCCCTCGGNGCCGCGTGTGCGGACCGAGCG 635

RESULT 11  
US-10-437-963-4649  
; Sequence 4649, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 4649  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_11513C.1  
US-10-437-963-4649

Query Match 22.9%; Score 308; DB 19; Length 696;  
Best Local Similarity 69.6%; Pred. No. 2.4e-77;  
Matches 451; Conservative 0; Mismatches 185; Indels 12; Gaps 2;  
QY 286 GCGCAGCCCGGGAGCGCATCTCACCCTCGTCCTCGACGACCTCTCGCGGGTTTTCAG 345  
DB 49 GCGTCGCGGACAGCTGGGCGCGTCTGACGCGGAGGGCTCTTCTACAGGCCCTC 108  
QY 346 GCGCCTCGTCGACAGATCAAGATCGCGCCATCATCGTGGCCACCGAGTTTCGCGGAG 405  
DB 109 CAGGCGGGAGCGCGGGAGCACAGAGCGGCTTTTACGCGCGTTTACCGCGCACCGC 168  
QY 406 CCCTACATCGCCAGTGCCTCGCCAAAGACCGCGGAGCACAGAGCGTTTCTGCTCGATT 465  
DB 169 GCGTCTCCGCGCGGTCCTCGCCAAAGACCGCGGAGCACAGCGCGCTCTCTCGGCTTC 228  
QY 466 GCGTCTCCGCGCTCCGAGTGTGTC---GGCCCGAGGCGCGCTGTGGCGGACGAGCGC 522  
DB 229 GCGTCTCCGCGCTCCGAGTGTGTCGATGCCCGGGGCGCGCTGTGGCGCGGACCGG 288  
QY 523 CGGAGGTGAAGGCGCATCGACACCGCCGCGCTCGCCCGCTGCTCCGCGCTACGTGCA 582  
DB 289 TCGAGCTGAAGGGGATCGACGCGCGGGTTCGCGCGCTGCTCCGCGCTACGTGTC 348  
QY 583 TCCGTTGCGGAGGAGTGGACTGTGCTGCGCTCGCGCGCGGCTTACGGAGGAAAGGT 642  
DB 349 ACGGCGCGCGCGCGGCTGACTGCGCTGCGCGCGCGGCTTACGGAGGAGGAGG 408  
QY 643 GGAGTCTTGTCACTCGCGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCAC 702  
DB 409 GCGTCTTGTGCTGAGTTCGCGGAGCTCAAGGCGTGTTCGAGGAGCAACCTGTACAC 468  
QY 703 TTCTACTCGCGCTCAATTTCTTGGGT-----ATGATGCTGCTGAGTCGAGCA 753  
DB 469 TTCTACTCGCGCTCAATTTCTTGGGT-----ATGATGCTGCTGAGTCGAGCA 528  
QY 754 GCGGAGATGGGGTGGGTGACGCTGAGTGGTGGGCTTTGCCCATGTGCGCGAGG 813  
DB 529 GGAGTGAAGCGCGGTGTAAAGGTGAAGTGTGGGACTTCGCGCATGTGACGATGG 588  
QY 814 GATGGGTGATTGACCACTTCTCGGCGGGCTCTGCTGCTGATCAAGTTCGTTTCT 873  
DB 589 GACGGGTGATTGACCACTTCTTGGCGGGCTCTGCTGCTCATCAAGTTTCATCGC 648  
QY 874 GACATTGTTCCGAGACTCTCATACGAGCGCTTTGGGTCTCTTTAA 921  
DB 649 GACATTGTTCCGAGACTCTCATACGAGCGCTTTTTCAGATCATTTCTTGA 696

RESULT 12  
US-10-042-894A-18  
; Sequence 18, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Caboon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof

```
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc feature
; LOCATION: (1)..(519)
; OTHER INFORMATION: n = A, T, C, or G
US-10-042-894A-18

Query Match      22.2%; Score 298.4; DB 14; Length 519;
Best Local Similarity 89.4%; Pred. No. 1.2e-74;
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 627 GTACGAGGAAAGGTGGAGTCTTGTACACAGCTGCGCGAGCTCAAGGGCTGGTTCGAGGA 686
Db 2 GTACGAGGAAAGGTGGAGTCTTGTACACAGCTGCGCGAGCTCAANGCGTGGTTCGAGGG 61

QY 687 GCAGACTCTGTCCACTTCTACTCGCGTCAATTTCTTCTGGGTATGATGCTGCTCAAGT 746
Db 62 GCAGACTCTGTCCACTTCTACTCGCGTCAATTTCTTCTGGGTATGATGCTGCTCAAGT 121

QY 747 CGCAGCAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGC 806
Db 122 CGCAGCAGCGGANGTGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGC 181

QY 807 CGAGGGTATGGGGTGAATGACACAACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTT 866
Db 182 CGAGGGTATGGGGTGAATGACACAACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTT 241

QY 867 CGTTCTGACATGTTCCGGAGTCTCTCATAGCAGCCCTTTGGGTCTTCTTAAGAGAG 926
Db 242 TGTTCTGACATGTTCCAGAGACTCTCTAGACGACGCTTTGGGTCTTCTTAAGAGAA 301

QY 927 GATCTCTGGCA-TTTCGATTGTAAACAAAGCCCTACAAGTTTTTGTCTGTGAAAGAGCG 985
Db 302 GATCTCTGGCA-TTTCGATTGTAAAGGAAACAAAGGAAACACTTTCAGCTGCCAAAGGAA 361

QY 986 CCTCCGA 992
Db 362 CCAGTGA 368

RESULT 13
US-10-425-115-10743
; Sequence 10743, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 10743
; LENGTH: 295
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: unsure
; LOCATION: (1)..(295)
; OTHER INFORMATION: unsure at all n locations

; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2002-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc feature
; LOCATION: (1)..(353)
; OTHER INFORMATION: n = A, T, C, or G
US-10-042-894A-19

Query Match      14.3%; Score 191.6; DB 14; Length 353;
Best Local Similarity 90.2%; Pred. No. 3.1e-44;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCAGGCGTGGTTCGAGGAGAGACTCTGTTCACCTTCTACTCGGGCTGATTTCTCTG 726
Db 1 CTCAGGCGTGGTTCGAGGAGAGACTCTGTTCACCTTCTACTCGGGCTGATTTCTCTG 60

QY 727 GGCTATGATGCTGCTGCTGAGGAGAGACTCTGTTCACCTTCTACTCGGGCTGATTTCTCTG 786
Db 61 GGCTATGATGCTGCTGCTGAGGAGAGACTCTGTTCACCTTCTACTCGGGCTGATTTCTCTG 120

QY 787 GTGGACTTTGCCCATGTGCGCGAGGGTATGGGGT--GATTGACCAACAACTTCTCTGGCG 844
Db 121 GTGGACTTTGCCCATGTGCGCGAGGGTATGGGGTGTGATTTGACCAACAACTTCTCTGGCG 180
```

QY 845 GGCTCTGCTCGCTGATCAAGTTC--GTTTCTGACATTGTTCCGGAGACTCCTCATACG-C 901  
Db 181 AGCTCTGCTAGCTGATCAAGTTTCCGGTTTCTTGACATGTTCCAGANACTCCTTAGACGCC 240  
QY 902 AGCCTTTGGGTCTCTCTTAAGAGA 925  
Db 241 AGCCTTTGGGTCTCTCTTAAGAAA 264

RESULT 15  
US-10-042-894A-11  
; Sequence 11, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1195  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (116)...(1048)  
US-10-042-894A-11

Query Match 13.7%; Score 184.6; DB 14; Length 1195;  
Best Local Similarity 55.6%; Pred. No. 4.2e-42;  
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY 56 CGGACCTCCACCGCGGAGCACCAGTCCGCGGCCACCGCCTCCGCGCAGCAAGCTGG 115  
Db 114 CATGCTCAAGTCCCGATCATCAAGTCCCGGTCCCGGAGACGGGGGAAAGCTGG 173  
QY 116 GCCCGCTCATCGACGCTCCGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGTGGGG 175  
Db 174 GGCCACTGGTGATGATTCCGGCGGCTTCTATAAGCCTCTCCAGAGCGATCATCGCGGAG 233  
QY 176 AGCAGAGGTCCGCTTCTATGAGCGTTCCTCCGCCACGCGCGCTCCCGGCCCGCATCC 235  
Db 234 ACACGGAAGTGGCTTTTACGAGTCAATCTATTCCAATACCGAGATCCCGGTCACATTC 293  
QY 236 GAGACACCTTCTCCCGGTTCCAGGCAACCGGACTCCTCCCGGAGGCGGAGCCCG 295  
Db 294 G---CAAAATCTCTCTGCTTTCAGGAACTAAGACTAT---TGAGGCGTCTGATGGAT 347  
QY 296 GGGAGCGCGATCCTCACCTCGTCTCGACGACTCTCCCGGGGTTTCAGGCGCCCTGG 355  
Db 348 CGGGTCTCAACCTCACTGTTCTGGAGATCTCGTCTCGGTCGCACGAACCCATCTC 407  
QY 356 TCGCAGACATCAAGATCGCGGCATCACTGGCCACCGAGTTCCCGGAGCCCTACATCG 415  
Db 408 TCATGACATCAAGACTGGATCCAGAACATGGTATCCGGAGGCTCTGAGGAGTACATCC 467  
QY 416 CCAAGTCCCTCCCAAGGACCGGGACCGAGCGTCTGCTCGGATTCGCGTCCG 475  
Db 468 AAAAGTCTTAGAGAAAGATCCAAATAGCACAAAGCGTTTCATTGGGTTTAGGATTTCTG 527  
QY 476 GCGTCCGAGTCTCGG---CCCGAGGGCGCGGTGTGGCGGACGGAGCGCCCGGAGGTGA 532  
Db 528 GCCTAAGGGTATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAGAGTTGTTT 587

QY 533 AGCCCATGGACACCCCGCGCTCCCGCGGTGCTCCGGCGGTACGTGTCTATCC----- 585  
Db 588 ATAGCTTTAATGCGGACGGTGTCAAGTCCGGCTCTCAGGAAGTTTGTCTTCCAACTTGT 647  
QY 586 -----GTTGCCGACGAGGGATGGACTGTGCGCTCCCGCGCGGTGTACGGAGGAAAAAG 640  
Db 648 CTCTGGGTCCAAATGTGGATCCGGATTGTTGTATGCAATCAAAAGTTTACTGTCAACGG 707  
QY 641 GTGGAGTCTTGTACAGCTGCGCGAGCTCAAGGCGTGGTTCGAGGAGCAGACTCTGTTC 700  
Db 708 GTGGAATTTTGGCACAATTGCTTCAGCTGAAGGAATGGTTTGAGGTTCAGACGAATATC 767  
QY 701 ACTTCTACTCGGCTCGGATTTCTTGGGCTATGATGCTGCTGAGTCCGAGCAGGCGGAG 760  
Db 768 ACTTCTATTTCTTGTTCACCTATATCTTATATGACAGGGAGTCTGCTTTGGACGGC- 823  
QY 761 ATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTCGCCGAGGGGTGATGGG 820  
Db 824 --TGTGCACACCCGAAAGTTAACTGGTGGACTTTGCACATGTGATGGATGCCACGGCG 881  
QY 821 TGATTGACCAACAATTCCTCGGCGGGCTCTGCTCGCTGATCAAGTTTCGTTTCTGACATTG 880  
Db 882 TGATCGATCACAACCTTCTTGGTGGCCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941  
QY 881 TTCCGGA 887  
Db 942 CTGATGA 948

Search completed: June 17, 2005, 04:00:54  
Job time : 893 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 22:26:07 ; Search time 4769 Seconds  
(without alignments)  
10727.276 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hic:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gesi:.\*  
9: gb\_gesi2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	921	68.5	960	CC344833	CG1A020TV
2	839.4	62.5	923	CG240397	OGYIC27TV
3	762.2	56.7	818	CG453578	OG9AD40TV
4	752	56.0	873	CC344824	CG1A020TH
5	644.8	48.0	867	CG724960	OGIAB23TV
6	620	46.1	801	CA134480	SCJFRT106
7	613.2	45.6	719	CG284572	OGYIF61TH
8	589.4	43.9	591	CG776236	1123005E0
9	564.6	42.0	656	CA245555	SCBGLF508
10	559	41.6	722	CA130685	SCCCRT100
11	557.8	41.5	3374	AY109355	Zea mays
12	555	41.3	671	CA202412	SCRLFL100
13	540.8	40.2	722	CG724950	OGIAB23TH
14	535.6	39.9	593	BZ774817	1151a11.b
15	525.6	39.1	744	CA130686	SCCCRT100
16	510.8	38.0	960	CG073495	PJUCG54TD
17	500	37.2	907	CG240387	OGYIC127TH
18	485	36.1	955	CG384766	PUHOG53TD
19	440.4	32.8	1131	CK211413	FGAS02325
20	435.6	32.4	794	BF617713	HVSMEO001
21	426.2	31.7	575	CG734025	1119161G0
22	423	31.5	551	BH408067	1007056A0
23	420.4	31.3	619	CA204213	SCAGFL109
24	376.2	28.0	842	BZ736665	OGEB508TM

25	376.2	28.0	866	9	CG367966	CG367966	OG0FJ44TH
26	376.2	28.0	939	9	CG367979	CG367979	OG0FJ44TV
27	372.8	27.7	588	5	BU999708	BU999708	H115H04r
28	351	26.1	927	9	CG284580	OGWIF61TV	CG284580
29	339.2	25.2	876	9	CG073494	PJUCG54TB	CG073494
30	334.2	24.9	431	9	CG776505	1123005E0	CG776505
31	334.2	24.9	557	2	BF277833	GA_EB003	BF277833
32	333.6	24.8	448	8	CC028437	3591_1_10	CC028437
33	332.8	24.8	1099	7	CK215908	FGAS02788	CK215908
34	324.8	24.2	1123	7	CK214494	FGAS02642	CK214494
35	319	23.7	453	4	BI956405	HVSMEN000	BI956405
36	311.2	23.2	433	4	BM099879	EBES01_SQ	BM099879
37	310.4	23.1	317	9	CG194403	FUIGW38TB	CG194403
38	309.8	23.1	598	9	CG286200	OGWAV62TV	CG286200
39	298	22.2	875	9	CC673013	OGWAV62TV	CG673013
40	297.2	22.1	620	5	BQ620191	TaLr1176F	BQ620191
41	293.6	21.8	409	4	BM099878	EBES01_SQ	BM099878
42	288.8	21.5	478	4	BG417083	HVSMER001	BG417083
43	277	20.6	306	8	CC023125	3591_1_33	CC023125
44	267.2	19.9	606	5	BU051170	111039E0	BU051170
45	250.4	18.6	506	9	CG366904	OG0AZ50TH	CG366904

#### ALIGNMENTS

RESULT 1  
CC344833  
LOCUS CC344833 960 bp DNA linear GSS 16-MAY-2003  
DEFINITION OG1A020TV ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0359C15,  
genomic survey sequence.  
ACCESSION CC344833  
VERSION CC344833.1 GI:30814239  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 960)  
AUTHORS Whiteley,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Contact: Cathy Whitelaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1..960  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBMA0359C15"  
/clone\_lib="ZM 0.7\_1.5\_KB"  
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

Query Match 68.5%; Score 921; DB 8; Length 960;  
Best Local Similarity 100.0%; Pred. No. 9.1e-206;  
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GTCCCTTCCCATACATGTCGAGCTCCACCGCGGAGACCAAGTCGCGGCCACC 94

Db 1 GTCCCTTCCCATACATGTCGAGCTCCACCGCGGAGACCAAGTCGCGGCCACC 60

QY	95	GGCGCTCGCCAGCAAGCTGGCCCGCTCATCGAGGCTCGCGCTCTTCTTCAAGCCGC	154
Db	61	GGCGCTCGCCAGCAAGCTGGCCCGCTCATCGAGGCTCGCGCTCTTCTTCAAGCCGC	120
QY	155	TCAGGCGCGCCAGCGTGGGAGCAGAGTGCCTTCTATGAGGCTTCTCCGCCACG	214
Db	121	TCAGGCGCGCCAGCGTGGGAGCAGAGTGCCTTCTATGAGGCTTCTCCGCCACG	180
QY	215	CGCGCTCGCGCCAGCTCGAGACACCTTCTTCCCGGTTCCAGGCACGCACTCC	274
Db	181	CGCGCTCGCGCCAGCTCGAGACACCTTCTTCCCGGTTCCAGGCACGCACTCC	240
QY	275	TCCACAGAGCGAGCGCGGAGCGCAGCTCCTCAGCTCGCTCGAGACCTCCTCG	334
Db	241	TCCACAGAGCGAGCGCGGAGCGCAGCTCCTCAGCTCGCTCGAGACCTCCTCG	300
QY	335	CGGGGTTTCAGGCGCCCTCGTTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGA	394
Db	301	CGGGGTTTCAGGCGCCCTCGTTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGA	360
QY	395	GTTCCGCGAGCCCTACATCGCAAGTCTCGCCAGAGACCGCGGACACAGAGGTT	454
Db	361	GTTCCGCGAGCCCTACATCGCAAGTCTCGCCAGAGACCGCGGACACAGAGGTT	420
QY	455	TGCTCGATTTCGCGTCTCCGCTCCGAGTCTGTCGGCCCGAGGCGCGGTGTGCGGA	514
Db	421	TGCTCGATTTCGCGTCTCCGCTCCGAGTCTGTCGGCCCGAGGCGCGGTGTGCGGA	480
QY	515	CGAGGCGCGGAGGTGAAGGCCATGAGACACCGCGCGGTTCGCGCGGTCTCCGCGCT	574
Db	481	CGAGGCGCGGAGGTGAAGGCCATGAGACACCGCGCGGTTCGCGCGGTCTCCGCGCT	540
QY	575	AGTGTATTCGTTGCCGAGAGGGATGAGCTGTGCGTTCGCGCGCGGTGTACGGAG	634
Db	541	AGTGTATTCGTTGCCGAGAGGGATGAGCTGTGCGTTCGCGCGCGGTGTACGGAG	600
QY	635	GAAAGGTGGAGTCTGTACAGCTGCGAGCTCAAGCGTGAAGCGTGTTCGAGGAGCAGACTC	694
Db	601	GAAAGGTGGAGTCTGTACAGCTGCGAGCTCAAGCGTGAAGCGTGTTCGAGGAGCAGACTC	660
QY	695	TGTTCCACTTCTACTCGCGGTGATCTTCTTGGGCTATGATGTGTCGAGTCGAGCAG	754
Db	661	TGTTCCACTTCTACTCGCGGTGATCTTCTTGGGCTATGATGTGTCGAGTCGAGCAG	720
QY	755	CGGAGATGGGGTGGGTGAGCGTGAAGCTGTGAGCTTTGCCATGTGGCCGAGGGTG	814
Db	721	CGGAGATGGGGTGGGTGAGCGTGAAGCTGTGAGCTTTGCCATGTGGCCGAGGGTG	780
QY	815	ATGGGTGATTGACCAACTTCTCGGGCGGCTCTGCTCGTGTATCAAGTTCTGTTCTG	874
Db	781	ATGGGTGATTGACCAACTTCTCGGGCGGCTCTGCTCGTGTATCAAGTTCTGTTCTG	840
QY	875	ACATTGTTCCGGAGACTCTCTATCAGCAGCTTTTGGGTCTCTTTAAAGAGAGATCTCG	934
Db	841	ACATTGTTCCGGAGACTCTCTATCAGCAGCTTTTGGGTCTCTTTAAAGAGAGATCTCG	900
QY	935	CATTTCGATTGATAACAAAG	955
Db	901	CATTTCGATTGATAACAAAG	921

RESULT 2  
CG240397  
LOCUS  
DEFINITION  
CG240397 923 bp DNA linear GSS 22-AUG-2003  
genomic survey sequence.  
ACCESSION  
CG240397  
VERSION  
CG240397.1 GI:34140283  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
Zea mays  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

REFERENCE AUTHORS	clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 923) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Other GSSs: OGYCI27TH Contact: Cathy Whitelaw TIGR
TITLE JOURNAL COMMENT	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TP Class: sheared ends.
FEATURES source	Location/Qualifiers 1..923 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBMA0788E05" /clone_lib="ZM 0.7 1.5_KB" /note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN	
Query Match	62.5%; Score 839.4; DB 9; Length 923;
Best Local Similarity	97.3%; Pred. No. 1.4e-186;
Matches	875; Conservative 0; Mismatches 21; Indels 3; Gaps 2;
QY	42 TCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCGCCACCGCGCCTC 101
Db	24 TCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCTCA 83
QY	102 CGCCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTACAAGCCGCTCCAGGC 161
Db	84 CGCCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTACAGCCGCTCCAGGC 143
QY	162 CGCGACCGTGGGAGCAGAGTCCCTTCTATGAGGCTTCTCGGCCACCGCGCGCT 221
Db	144 CGCGACCGTGGGAGCAGAGTCCCTTCTATGAGGCTTCTCGGCCACCGCGCGCT 203
QY	222 CCCGCCGCATCCGAGACACCTTCTCCCGGTTCCAGCGACGCGACTCTCTCCCCAC 281
Db	204 CCCGCCGCATCCGAGACACCTTCTCCCGGTTCCAGCGACGCGACTCTCTCTCCAC 263
QY	282 CGAGCGCAGCCCGGGAGCGCATCTCACCTCGTCTCGACGACCTCTCGCGGGTT 341
Db	264 CGAGCGCAGCCCGGGAGCGCATCCGCACCTCGTCTCGACGACCTCTCGCGGGTT 323
QY	342 TCAGCGCGCTCGTTCGAGACATCAAGATCGCGCCATCACTGGGCCACCGAGTTCGCC 401
Db	324 TCAGCGCGCTCGTTCGAGACATCAAGATCGCGCCATCACTGGGCCACCGAGTTCGCC 383
QY	402 GGAGCGCTACATCGCCAGTGCCTCGCCAGACCGCGGACCGAGGTTCTGCTCG 461
Db	384 GGAGCGCTACATCGCCAAAGTACCTCGCCAAAGACCGCGGACCGAGGTTCTGCTCG 443
QY	462 ATTTCGCGTCTCGCGCTCCGAGTGTGTCGGCCCGAGGCGCGCTGTGCGGACCGAGCG 521
Db	444 ATTTCGCGTCT--TGGTCCGAGTGTGTCGGCCCGAGGCGCGCTGTGCGGACCGAGCG 501
QY	522 CCCGAGGTGAAGCCATGGACACCGCGCGCTCGCGCGCGTCTCGCGCGCTTACGTGTC 581
Db	502 CCCGAGGTGAAGGCTATGGACACCGTCCGCGCTCGCGCGCTTACGTGTC 561
QY	582 ATCCG-TTCCCGACGAGGGATGCTGTGCGCTCGCGCGCGGTGTACGAGGAAAG 640
Db	562 ATCCGTTTCCCGACGAGGGATGGATGCGCTCGCGCGCGGTGTACGAGGAAAG 621
QY	641 GTGGAGTCTTGTACAGTTCGCGAGCTCAAGGCGTGGTTTCGAGGAGCAGACTCTGTTCC 700



```
Db      622 GTGGAGTCTGTACAGCTGGCGAGCTCAAGCATGGTTGGAGGACAGACTCTGTCC 681
Qy      701 ACTTCTACTCGCGGTGATTTCTTCTGGGCTATGATGCTGCTGAGTCGACGAGCGGAG 760
Db      682 ACTTCTACTCGCGGTGATTTCTTCTGGGCTATGATGCTGCTGAGTCGACGAGCGGAG 741
Qy      761 ATGGGGTGGGGTACGGTGAAGCTGAGTCTTCCCATGTCGCGGAGGGTATGGGG 820
Db      742 GTGGGGTGGGGTACAGTGAAGCTGTGAGCTTTTCCCATGTCGCGGAGGGTATGGGG 801
Qy      821 TGATTGACCACAACTTCTTCTGGGCGGCTCTGCTGCTGATCAAGTTCTGTTTCTGACATTG 880
Db      802 TGATTGACCACAACTTCTTCTGGGCGAGCTCTGCTAGCTGATCAAGTTCTGTTTCTGACATTG 861
Qy      881 TTCGGAGACTCTCATAGCAGCCTTTGGGCTCTTCTTAAGAGAGATCTCGGCATTT 939
Db      862 TTCAGAGACTCTTAGAGCAGCCTTTGGGCTCTTCTTAAGAGAGATCTCGGCATTT 920

RESULT 3
CG453578      818 bp      DNA      linear      GSS 17-SBP-2003
LOCUS      OG9AD40TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0849H08,
DEFINITION      genomic survey sequence.
ACCESSION      CG453578
VERSION      CG453578.1 GI:34838578
KEYWORDS      GSS.
SOURCE      Zea mays
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 818)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG9AD40TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source      Location/Qualifiers
1..818
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0849H08"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      56.78; Score 762.2; DB 9; Length 818;
Best Local Similarity 99.68; Pred. No. 2e-168;
Matches 764; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy      8 GTGAGTCCGTCACCCCTCGGGCCCATAGTCCCTTCCCATACCATGTCGACCTCCACC 67
Db      52 GTGAGTCCGTCACCCCTCGGGCCCATAGTCCCTTCCCATACCATGTCGACCTCCACC 111
Qy      68 CGCGGAGACCAAGTCGCGGCCACCGCGCTTCCGCGCAGCAGCTGGGCGGCTCATCG 127
Db      112 CGCGGAGACCAAGTCGCGGCCACCGCGCTTCCGCGCAGCAGCTGGGCGGCTCATCG 171
Qy      128 ACGGCTCGGCGCTTCTTACAGCGGCTCCAGGCGCGGACCGTGGGGGACGAGGTCG 187
```

```
Db      172 ACGGCTCGGCGCTTCTTCTAAGCGCTCCAGGCGCGGACCGTGGGGAGCACAGGTCG 231
Qy      188 CTTTCTATAGAGGCTTCTCGGCCACGCGCGTCCGGCCCGGCATCCGAGACACCTTCT 247
Db      232 CTTTCTATAGAGGCTTCTCGGCCACGCGCGCTCCGGCCCGGCATCCGAGACACCTTCT 291
Qy      248 TCCCCCGGTTTCAACGGCAGCGACTCTCTCCACACCGAGGCGCAGCCGGGAGCCGCATC 307
Db      292 TCCCCCGGTTTCAACGGCAGCGACTCTCTCCACACCGAGGCGCAGCCGGGAGCCGCATC 351
Qy      308 CTCACCTCTCTCGACAGACTCTCTCGGGGGTTTCAGGGCCCTCGTCCGAGACATCA 367
Db      352 CTCACCTCTCTCGACAGACTCTCTCGGGGGTTTCAGGGCCCTCGTCCGAGACATCA 411
Qy      368 AGATCGGCGCCATCAGTGGCCACAGTTTCGCGGAGCCCTACATCGCCAAAGTGCCTCG 427
Db      412 AGATCGGCGCCATCAGTGGCCACAGTTTCGCGGAGCCCTACATCGCCAAAGTGCCTCG 471
Qy      428 CCAAGGACCGCGGACACAGAGCGTTTCTCGATTTCGCGCTCTCCGGCGTCCGAGTCG 487
Db      472 CCAAGGACCGCGGACACAGAGCGTTTCTCGATTTCGCGCTCTCCGGCGTCCGAGTCG 531
Qy      488 TCGGCCCCAGAGCGCGCTGTGCGGACCGAGCGCCCGGAGTGAAGGCCATGACACCG 547
Db      532 TCGGCCCCAGAGCGCGCTGTGCGGACCGAGCGCCCGGAGTGAAGGCCATGACACCG 591
Qy      548 CGGGCGTCCGCGCGCTGCTCCGGCGTACGTGTCATCCGTCGCGGAGGGATGAGCT 607
Db      592 CGGGCGTCCGCGCGCTGCTCCGGCGTACGTGTCATCCGTCGCGGAGGGATGAGCT 651
Qy      608 GTGCGCTCCCGCGCGGTGTACGAGAGAAAGGTGAGTCTTGTTCACAGCTCGCGGAGC 667
Db      652 GTGCGCTCCCGCGCGGTGTACGAGAGAAAGGTGAGTCTTGTTCACAGCTCGCGGAGC 711
Qy      668 TCAAGCGGTGTTTCGAGGACAGACTCTGTTCACATTCTTACTCGGCGTCAATCTTCTGG 727
Db      712 TCAAGCGGTGTTTCGAGGACAGACTCTGTTCACATTCTTACTCGGCGTCAATCTTCTGG 771
Qy      728 GCTATGATCTGCTGCACTCGAGTCGCGAGCGGAGATGGGGTGGGGTGG 774
Db      772 GCTATGATCTGCTGCACTCGAGTCGCGAGCGGAGATGGGGTGGGGTGG 818

RESULT 4
CC344824/c
LOCUS      OG1AQ20TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0359C15,
DEFINITION      genomic survey sequence.
ACCESSION      CC344824
VERSION      CC344824.1 GI:30814230
KEYWORDS      GSS.
SOURCE      Zea mays
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 873)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source      Location/Qualifiers
1..873
```

```
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0359C15"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBOSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      56.0%; Score 752; DB 8; Length 873;
Best Local Similarity 100.0%; Pred. No. 5.1e-166;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 CTCGCGCCACGCGCGCTCCGGCCGCGATCCGAGACACTTCTTCCCGCGGTTCACGG 263
DB 873 CTCGCGCCACGCGCGCTCCGGCCGCGATCCGAGACACTTCTTCCCGCGGTTCACGG 814
QY 264 CACGCGACTCTCTCCACACGAGCGCAGCCCGGGAGCGCATCTCTCACTCGTCTCGA 323
DB 813 CACGCGACTCTCTCCACACGAGCGCAGCCCGGGAGCGCATCTCTCACTCGTCTCGA 754
QY 324 CGACTCTCTCGGGGGTTTTCAGCGCCCTCGCTCGCAGACATCAAGATCGGGCCATCAC 383
DB 753 CGACTCTCTCGGGGGTTTTCAGCGCCCTCGCTCGCAGACATCAAGATCGGGCCATCAC 694
QY 384 GTGGCCACCGAGTTCCGCGGAGCCCTACATCGGCAAGTCCCTCGCCAAAGACCGGGGAC 443
DB 693 GTGGCCACCGAGTTCCGCGGAGCCCTACATCGCCTCAAGTCCCTCGCCAAAGACCGGGGAC 634
QY 444 CACGACGCTTCTGCTCGGATTCGCGTCTCCGGCTCCGAGTCCGAGTCCGCGCCGAGGGCGC 503
DB 633 CACGACGCTTCTGCTCGGATTCGCGTCTCCGGCTCCGAGTCCGCGCCGAGGGCGC 574
QY 504 CFTGTGGCGGACGAGCGCCGAGAGTGAAGGCATGACACCGCGCGGTCCGCGCGCT 563
DB 573 CFTGTGGCGGACGAGCGCCGAGAGTGAAGGCATGAGACCGCGCGGTCCGCGCGCT 514
QY 564 GCTCGGCGCTACGTGTCTATCGTTGCCGACGAGGGGATGACTGTGCGCTCGCGCGGC 623
DB 513 GCTCGGCGCTACGTGTCTATCGTTGCCGACGAGGGGATGACTGTGCGCTCGCGCGGC 454
QY 624 GTGTACGAGGAAAGGTGGAGTCTGTTCACAGCTCGCGAGCTCAAGCGGTGGTTGA 683
DB 453 GTGTACGAGGAAAGGTGGAGTCTGTTCACAGCTCGCGAGCTCAAGCGGTGGTTGA 394
QY 684 GGAGCAGACTCTGTTCACCTTCTACTCGGCGTCGATTTCTTCTGGGCTATGATGCTGCTGC 743
DB 393 GGAGCAGACTCTGTTCACCTTCTACTCGGCGTCGATTTCTTCTGGGCTATGATGCTGCTGC 334
QY 744 AGTCGACGACGCGGAGATGGGGGTGGGTGACCGGTGAAGCTGGTGGACTTTGCCCATGT 803
DB 333 AGTCGACGACGCGGAGATGGGGGTGGGTGACCGGTGAAGCTGGTGGACTTTGCCCATGT 274
QY 804 GSCGAGGGTGATGGGTGATTTGACACAACTTCTTGGCGGGCTCTGCTGCTCATCAA 863
DB 273 GSCGAGGGTGATGGGTGATTTGACACAACTTCTTGGCGGGCTCTGCTGCTCATCAA 214
QY 864 GTTCTGTTCTGACATTTGTTCCGAGACTCTCTCATACGACGCTTTGGGTCTCTTAAGA 923
DB 213 GTTCTGTTCTGACATTTGTTCCGAGACTCTCTCATACGACGCTTTGGGTCTCTTAAGA 154
QY 924 GAGGATCTCGGATTTGATTTGATAACAAAG 955
DB 153 GAGGATCTCGGATTTGATTTGATAACAAAG 122

RESULT 5
CC724960
LOCUS
DEFINITION
OGLAB23rtv ZM_0.7.1.5_KB Zea mays genomic clone ZMMBMA0307C22,
genomic survey sequence.
CC724960
ACCESSION
```

```
VERSION      CC724960.1  GI:32143893
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 867)
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
              Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
              Consortium for Maize Genomics
              Unpublished (2002)
              Other_GSSs: OGLAB23TH
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: tp
              Class: sheared ends.
FEATURES     Location/Qualifiers
              1..867
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone="ZMMBMA0307C22"
               /clone_lib="ZM 0.7 1.5 KB"
               /note="Vector: pBOSK-; Site 1: HincII; 0.7-1.5 kb
               methylation filtered genomic DNA library"
ORIGIN
Query Match      48.0%; Score 644.8; DB 9; Length 867;
Best Local Similarity 89.4%; Pred. No. 8.4e-141;
Matches 753; Conservative 0; Mismatches 27; Indels 62; Gaps 3;

QY 2 CACGAGGTCTAGTCCGTACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGACC 61
DB 87 CACCATCGCCACCCCTCACCCCTTCTCTCCATAG-----TCCCATACCATGTCGACC 140
QY 62 TCCACCCCGCGGAGCACCAAGTCGCGCGCCACCGCGCTCGCCACGAGCTGGGGCCGC 121
DB 141 TCCACCCCGCGGAGCACCAAGTCGCGCGCTCACCGCGCTTCGCCAGCAAGCTGGGCCAC 200
QY 122 TCATCGAGCGGTCCGGCCCTCTTCTACAAGCCGCTCCAGCGCGGACCGTGGGGAGCAG 181
DB 201 TCATCGAGCGACTTGGCCCTTCTTACAAGCCGCTCCAGCGCGGACCGTGGGGAGCAG 260
QY 182 AGTTCGCTTCTATGAGCGGTTCTCCGCCACCGCGCGTCCCGCCCGCATCCGAGACA 241
DB 261 AGTTCGCTTCTATGAGCGGTTCTCCGCCACCGCGCGTCCCGCCCGCATCCGAGACA 320
QY 242 CTTCTTCCCGCGTTCCACGCGACGCTCTCTCCACCGAGCGGACCGCGGGAGC 301
DB 321 CTTCTTCCCGCGTTCCACGCGACGCTCTCTCCACCGAGCGGACCGCGGGAGC 380
QY 302 CGCATCTCTCAGCTCGTCTCTCGACGACTCTCTCGGGGGTTTCAGCGCGCTTCGCTCGCAG 361
DB 381 CGCATCTCGACCTCGTCTCTCGACGACTCTCTCGGGGGTTTCAGCGCGCTTCGCTCGCAG 440
QY 362 ACATCAAGATCGGCGCATCATCGTGGCCACCGAGTTCCCGGAGCCCTACATCGCAAGT 421
DB 441 ACATCAAGATCGTGGCATTCAGT-----GACACGAGCGTTCTGCTCGGATTCGCGGTCC 464
QY 422 GCCTCGCAAGACCGCGGGACCAAGCGGTTCTGCTCGGATTCGCGGTTCGCGGTCC 481
DB 465 -----GACACGAGCGTTCTGCTCGGATTCGCGGTTCGCGGTCC 505
QY 482 GAGTCTCGGCGCGGAGGCGCGGTGTGGCGGACGAGCGCGCGGAGGTGAAGGCATGG 541
DB 506 GAGTCTCGGCGCGGAGGCGCGGTGTGGCGGACGAGCGCGCGGAGGTGAAGGCTATGG 565
```

542 ACACCGCGCGTCCGCGCGTCTCCGCGCTACGTGTCATCG-TTGCAGAGGGG 600  
 566 ACATTGTCGCGTCCGCGCGTCTCCGCGCTACGTGTCATCGTTCGCGAGGGG 625  
 601 ATGAGCTGTGCGTCCGCGCGGTGTACGAGGAAAGGTGAGTCTTTGTACAGCTG 660  
 626 ATGAGCTGTGCGTCCGCGCGGTGTACGAGGAAAGGTGAGTCTTTGTACAGCTG 685  
 661 CGGAGCTCAAGCGTGTGCGAGGACAGACTCTGTTCACTTCTACTCGCGTCAATT 720  
 686 CGGAGCTCAAGCGTGTGCGAGGACAGACTCTGTTCACTTCTACTCGCGTCAATT 745  
 721 CTTCTGGGCTATGATGCTGTCAGTGCAGTGCAGCGCGGAGATGGGGTGGCGTG 780  
 746 CTTCTGGGCTATGATGCTGTCAGTGCAGCGCGGAGATGGGGTGGCGTAAAGTG 805  
 781 AAGTGTGTGACTTTGCCCCATGTGCGCGAGGGTGATGGGGTGATGACCAACTTCTG 840  
 806 AAGTGTGTGACTTTGCCCCATGTGCGCGAGGGTGATGGGGTGATGACCAACTTCTG 865  
 841 GG 842  
 866 GG 867

RESULT 6  
 CAL34480  
 LOCUS  
 DEFINITION SCJFRT1061H11.9 RT1 Saccharum officinarum cDNA clone SCJFRT1061H11 5', mRNA sequence.  
 CAL34480  
 CAL34480.1 GI:35021536  
 EST.  
 Saccharum officinarum  
 Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.  
 1 (bases 1 to 801)  
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@unicamp.br  
 Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>  
 Plate: 061 row: H column: 11  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1..801  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCJFRT1061H11"  
 /lab\_host="DH10B"  
 /clone\_lib="RT1"  
 /note="Organ: Root tips (0.3cm-long) from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Root tips (0.3cm-long) from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library

construction can be obtained at  
<http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Query Match 46.1%; Score 620; DB 6; Length 801;  
 Best Local Similarity 89.2%; Pred. No. 5.7e-135;  
 Matches 701; Conservative 0; Mismatches 82; Indels 3; Gaps 3;

QY 37 CCCTTCCCATATGTCGACCTCCACCCGCGGAGCACCAAGTGCAGCGCCACCGC 96  
 DB 16 CTTCTCGCCCTACCATGTCGACCTCCGCGCGCGGAGCACCAAGTGCAGCGCCACCGC 75  
 QY 97 GCCTCCGCGCAGCAAGCTGGCGCTCATCGAGCTCCGCGCTCTTCTAAGACCGCTC 156  
 DB 76 GCCTCCGCGCAACAAGCTGGCTCATCGAGCTCTGGCTCTTCTAAGACCGCTC 135  
 QY 157 CAGGCGCGGACCGCTGGGAGCAGAGTGGCTTCTATGAGGGCTTCTCGGCCAGCC 216  
 DB 136 CAGGCTGGGGACCGCGGGGAGCAGAGCTCGCTTCTAGAGGCGTCTCCACCGCC 195  
 QY 217 GCGTCCGCGCGCATCCGAGACACTTCTCCCGGTTCCACGCGACGCGACTCCTC 276  
 DB 196 GCGTCCGCGCGCATCCGAGACACTTCTCCCGGTTTCAAGGACGCGACTCCTC 255  
 QY 277 CCACCGAGCGCAGCGCGGAGCGCATCTCTACCTCTCTCGAGAGCTCTCTCGCG 336  
 DB 256 CCACCGAGCGCGCGGAGCGCATCTCTCGAGAGCTCTCTCGAGAGCTCTCTCGCG 315  
 QY 337 GGTTCAGCGCGCTGCTCGAGACATCAAGATCGGCGCCATCAC-GTGGCCACCGAG 395  
 DB 316 GGGCTGGAGCGCGCTCGCTCGCGGATCAAGATCGGCGCCATCAAGTGGCGCGCGAG 375  
 QY 396 TTGCG-CGAGAGCGCTACATCGCAAGTCCCTCCGAGGACCGCGGACCAAGCGCTTC 454  
 DB 376 CTGCGAGAGCGCTAGCTGCTCAAGTCCCTGTCAGAGACCGCGGACCAAGCGCTTC 435  
 QY 455 TGCTCGGATTCGCGCTCTCGCGCTCCGAGTCTCGCGCCCGAGGGCGCGTGTGGCGGA 514  
 DB 436 TGCTCGGATTCGCGCTCTCGCGCTCCGCGTCTCGGTCCGAGGGCGCGTGTGGCGGA 495  
 QY 515 CGAGCGCGCGAGGTGAAGGCCATGACACCGCGCGCTCCGCGCGCTCTCGCGCGCT 574  
 DB 496 CTGAGCGCGCGAGGTGAAGGCCCTCGACACCGCGCGCTCCGCGCGCTCTCGCGCGCT 555  
 QY 575 ACCTGTATCCGTTGCGAGCGGAGTGGACTGCGCTCGCGCGCGGTGTACCGAG 634  
 DB 556 ACCTGTATCCGTTGCGAGCGGAGTGGACTGCGCACTCGCGCGCGGTGTACCGAG 615  
 QY 635 GAAAGGTGAGTCTTGTCTACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGAGCAGACTC 694  
 DB 616 GGAAGGGAGTCTTGTCAATTTGGCGAGCTTAAGCGTGTGTTGAGGACAGACTC 675  
 QY 695 TGTTCCACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGCTGAGTGCAGCAG 754  
 DB 676 TGCTCCACCTTTACTCGGGGCGATTCTTCTGGGCTATGAAGCTAGTGTGCGCTTGCACAA 735  
 QY 755 GCGG-AGATGGGGTGGGTGACGCTGAAGCTTGTGAGCTTTGCCCATGTGGCGAGGGT 813  
 DB 736 CCGAAGGTAAAGGTGGGGTGAAGGTAAAGTGGGGCCCTTTGCCCATGTGGCCCAAGG 795  
 QY 814 GATGGG 819  
 DB 796 GAAGGG 801

RESULT 7  
 LOCUS  
 DEFINITION OSW161TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBma0584K02, genomic survey sequence.  
 ACCESSION  
 VERSION CG284572  
 KEYWORDS CG284572.1 GI:34198786 GSS.

SOURCE ORGANISM	Zea mays		Db			486	GCCGACGAGGGGATGGAGTCTGCGCGCTCGCGCGCGGTGTACGAGAGAAAAGGTGGAGTC		545
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.								
REFERENCE AUTHORS	1 (bases 1 to 719) Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N. Consortium for Maize Genomics Unpublished (2002) Other GSSs: OGWIF61TV Contact: Cathy Whitelaw		Qy			649	TTGTCACTAGCTCGCGAGCTCAAGCGGTGTTTCGAGGACGACTCTGTTCACCTTCTAC		708
	TIGR								
TITLE JOURNAL COMMENT	7912 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends.		Db			546	TTGTCACTAGCTCGCGAGCTCAAGCGGTGTTTCGAGGGGAGACTCTGTTCACCTTCTAC		605
	TIGR								
FEATURES source	Location/Qualifiers 1..719 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMBMA0594K02" /clone_lib="ZM_0.7_1.5_KB" /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"		Qy			709	TCGGCGTCGATTCTTCTCGGGCTATGCTGCTGCAGTCGCGAGCGCGAGATGGGGGT		768
	TIGR								
ORIGIN	45.6%; Score 613.2; DB 9; Length 719; Query Match 90.4%; Pred. No. 2.2e-133; Matches 700; Conservative 0; Mismatches 18; Indels 56; Gaps 2;		Db			769	GGGGTGACGGTGAAGCTGTGGAGCTTTGCCCATGTGGCCGAGGGTGCATGGGGTG		822
	50 CCATGTCGACCTCCACCGCGGAGCACCAGTCGCGGCCACCGCGCTCGCGCAGCA								
FEATURES source	Location/Qualifiers 1..591 /organism="Zea mays" /mol_type="genomic DNA" /cultivar="mixed background W23/A188/B73/K55" /db_xref="taxon:4577" /tissue_type="leaf" /dev stage="adult" /lab_host="DH10B" /clone_lib="1123 - RescueMu Grid L" /note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid L was grown in Molokai in 2001. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."		Qy			606	TCGGCGTCGATTCTTCTCGGGCTATGCTGCTGCAGTCGCGAGCGCGAGGTGGGGT		665
	1123005E08.xl 1123 - RescueMu Grid L Zea mays genomic, genomic survey sequence.								
ORIGIN	43.9%; Score 589.4; DB 9; Length 591; Query Match 99.8%; Pred. No. 8.7e-128; Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		Db			25	CGCGCCCATAGTCCCTTCCCATACCATGTCCGACCTCCACCGCGGAGCACCAGTTC		84
	1 CGCGCCCATAGTCCCTTCCCATACCATGTCCGACCTCCACCGCGGAGCACCAGTTC								

```
QY 85 GCCGGCCACCGCGCTCCGCGCAGCAGCTGGGCGCGCTCATCGACGGCTCCGGCCCTCTTC 144
Db 61 GCCGGCCACCGCGCTCCGCGCAGCAGCTGGGCGCGCTCATCGACGGCTCCGGCCCTCTTC 120
QY 145 TACAAGCCGCTCCAGGCGCGCGCAGCAGCTGGGAGCAGAGGTGCGCTTCTATGAGGGCTTC 204
Db 121 TACAAGCCGCTCCAGGCGCGCGCAGCAGCTGGGAGCAGAGGTGCGCTTCTATGAGGGCTTC 180
QY 205 TCCGCCACCGCGCGCTCCGCGCGCGCATCCGAGACACCTTCTTCCCGCGGTTCACGGC 264
Db 181 TCCGCCACCGCGCGCTCCGCGCGCGCATCCGAGACACCTTCTTCCCGCGGTTCACGGC 240
QY 265 ACCGCGACTCTCCCGCAGGCGCGCAGCGCGGAGCGCGCATCTCACCCTGCTCTCGAC 324
Db 241 ACCGCGACTCTCCCGCAGGCGCGCAGCGCGGAGCGCGCATCTCACCCTGCTCTCGAC 300
QY 325 GACTCTCTCCGCGGGTTTCAAGGGCGCTTCCGCTCGCAGACATCAAGATCGCGCGCATCACG 384
Db 301 GACTCTCTCCGCGGGTTTCAAGGGCGCTTCCGCTCGCAGACATCAAGATCGCGCGCATCACG 360
QY 385 TGCCACCGAGTTCCGCGGAGCCTACATCCCAAGTGCCTCGCAGAGACCGCGGACC 444
Db 361 TGCCACCGAGTTCCGCGGAGCCTACATCCCAAGTGCCTCGCAGAGACCGCGGACC 420
QY 445 ACAGCGCTTCTGCTCGGATTCGCGTCTCCGGGTCCGAGTCTCGCGCCCGAGGGGCC 504
Db 421 ACAGCGCTTCTGCTCGGATTCGCGTCTCCGGGTCCGAGTCTCGCGCCCGAGGGGCC 480
QY 505 GTGTGGCGGACGAGCGCCCGAGGTGAAGCCATGAGACACCGCGCGCTCCGCGCGTG 564
Db 481 GTGTGGCGGACGAGCGCCCGAGGTGAAGCCATGAGACACCGCGCGCTCCGCGCGTG 540
QY 565 CTCGGCGCTACGTGTTCATCCGTTGCGGACGAGGGATGACGTGTGCGCTC 615
Db 541 CTCGGCGCTACGTGTTCATCCGTTGCGGACGAGGGATGACGTGTGCGCTC 591
```

## RESULT 9

CA245555  
LOCUS  
DEFINITION  
SCBGF5081A03.9 Saccharum officinarum FL5 Saccharum officinarum  
cDNA clone SCBGF5081A03 5', mRNA sequence.

CA245555.1 GI:35324301

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Saccharum officinarum  
Saccharum officinarum  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 656)  
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenhariaia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089

Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 081 row: A column: 03

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. 656

/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

## FEATURES

## Source

/clone="SCBGF5081A03"

/lab\_host="DH10B"

/clone\_lib="Saccharum officinarum FL5"

/notes="Organ: Developed inflorescence (20cm-long) without

rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [Developed

inflorescence (20cm-long) without rachis]. cDNA was

prepared from polyA+ mRNA using SuperScript plasmid

System Kit (Invitrogen). The double-strand cDNAs were

fractionated in a sepharose CL-2B 40cm-columns and

fragments sizing between 0.8 and 1.5 kb were

directionally cloned into the vector. Details of each

source of RNA and library construction can be obtained at

http://sucet.lad.ic.unicamp.br/public"

## ORIGIN

Query Match 42.0%; Score 564.6; DB 6; Length 656;  
Best Local Similarity 93.7%; Pred. No. 6e-122;  
Matches 599; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 37 CCCTTCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGC 96

Db 18 CCTCTCGCCCTCACCATGTCGACCTCCGCGCGGAGCACCAAGTCGCGGCCACCGC 77

QY 97 GCCTCCGCGAGCAAGCTGGCGCGCTCATCGAGCGCTCCGCGCTCTTCTACAGCGCGCTC 156

Db 78 GCTCTCGCCCAACAGCTGGGTCCGCTCATCGAGCGCTCTGGCCCTCTTCTACAGCGCGCTC 137

QY 157 CAGCGCGCGGACCGTGGGAGCAGCAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGC 216

Db 138 CAGGTCCGGGACCGCGCGGAGCAGCAGGTGCGCTTCTACAGCGCTTCTCCGCCACCGC 197

QY 217 GCGTCCCGCGCGATCCGAGACACCTTCTCCCGCGTTCACCGGACCGGACTCTCTC 276

Db 198 GCGTCCCGCGCGATCCGAGACACCTTCTCCCGCGTTCACCGGACCGGACTCTCTC 257

QY 277 CCACCGAGCGCAGCGCGGAGCGGATCTCTCAGCTCTGCTCTCGACGACCTCTCTCGCG 336

Db 258 CCCACCGAGCGCGGAGCGGAGCGGATCTCTCAGCTCTGCTCTCGACGACCTCTCTCGCG 317

QY 337 GGCTTTACGGCGCTTCTGCTGCGAGACATCAAGATCGCGGCCATCAGTGGCCACCGAGT 396

Db 318 GGCTGGAGCGCGCTTCTGCTGCGAGACATCAAGATCGCGGCCATCAGTGGCCACCGAGT 377

QY 397 TCGCGGAGCGCTTCTGCTGCGAGTCTCGCGCGCGGAGCGGAGCGGAGCGGAGCGGAG 456

Db 378 TCGCGGAGCGCTTCTGCTGCGAGTCTCGCGCGCGGAGCGGAGCGGAGCGGAGCGGAG 437

QY 457 CTGCGATTCTCGCGCTTCTGCTGCGAGTCTCGCGCGCGGAGCGGAGCGGAGCGGAGCGGAG 516

Db 438 CTGCGATTCTCGCGCTTCTGCTGCGAGTCTCGCGCGCGGAGCGGAGCGGAGCGGAGCGGAG 497

QY 517 GAGCGCGCGGAGGTGAAGGCCATCGACCGCGCGGCTTCTCGCGCGGCTTCTCGCGCGGCTTCT 576

Db 498 GAGCGCGCGGAGGTGAAGGCCATCGACCGCGCGGCTTCTCGCGCGGCTTCTCGCGCGGCTTCT 557

QY 577 GTGTCTATCTGCTTCCCGACGAGGGATGAGTGTGCGCTCGCGCGGCGGCTTCTCGCGCGGCTTCT 635

Db 558 GTTTCATCTGCTTCCCGACGAGGGATGAGTGTGCGCTCGCGCGGCGGCTTCTCGCGCGGCTTCT 617

QY 636 AAAAGGTGGAGTCTTCTGTCACAGCTGCGCGGAGCTCAAGGC 674

Db 618 AAAAGGTGGAGTCTTCTGTCACAGCTGCGCGGAGCTTAAAGC 656

## RESULT 10

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

CA130685 722 bp mRNA linear EST 24-SEP-2003

SCCRR1004G05.g RT1 Saccharum officinarum cDNA clone SCCRR1004G05

5', mRNA sequence.

CA130685

CA130685.1 GI:35013914

EST.

Db	375	TCGCCGAGCCCTACGTCGCCAAGTCTCGCCAAAGGACCGCGGACCACGAGCATCTCTG	434
Qy	457	CTCGATTTCGGGTCTCCGGCGTCCGAGTCTGTCGCCCGGAGGCGCGTGTGGCGGACG	516
Db	435	CTCGATTTCGGGTCTCCGGGTCTCCGGGTCTCGGTCCCGAGGGGCGCGTGTGGCGGACT	494
Qy	517	GAGGCCCGGAGGTGAAGGCCATGGACACCGCCGGCGTCCGCCCGGTGCTCCGGCGCTAC	576
Db	495	GAGGCCCGGAGGTGAAGGCCCTGGACACCGCCGGCGTCCGCCCGGTGCTCCGGCGCTTA	554
Qy	577	GTGTCATCCGTTGCCGAGCGGAGTGACTGTGCGTTCGCCCGCGGTGTA-CGGAGG	635
Db	555	CGTTTCATCCGTTGCCGAGCGGAGTGACTGTGCGTTCGCCCGCGGTGTA-CGGAGG	614
Qy	636	AAAAGGTGGAGTCTTGTTCACAGCTCGCGAGCTCAAGCGGTGGTTTCGAGGAGCAGACTCT	695
Db	615	CAAAAGGGAGTCTTGTTCACAGCTCGCGAGCTTAAAGCGTGGTTCGAGGA-CAAACTCT	673
Qy	696	GTTCACACTTCTACTCGCGTTCGATTTCTTCTGGGCTATGATGCT	738
Db	674	GGATCACATT-TACTCGGGCCGATTTCTTTCGGGCTAATATGCT	715
RESULT 11			
AY109355			
LOCUS	AY109355	3374 bp	mRNA linear HTC 17-OCT-2002
DEFINITION	Zea mays CL49_2 mRNA sequence..		
ACCESSION	AY109355		
VERSION	AY109355.1	GI:21213015	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 3374) Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 3374)		
AUTHORS	Coe, E.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.		
FEATURES	Location/Qualifiers		
source	1. .3374 /organism="Zea mays" /mol_type="mRNA" /db_xref="MaizeDB:632179" /db_xref="taxon:4577" /clone_lib="Maize Mapping Project/DuPont Consensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
ORIGIN			
Query Match	41.5%;	Score 557.8;	DB 3; Length 3374;
Best Local Similarity	67.7%;	Pred. No. 2.8e-120;	
Matches	672; Conservative	0; Mismatches 259;	Indels 62; Gaps 3;

SOURCE	Saccharum officinarum		
ORGANISM	Saccharum officinarum		
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.		
REFERENCE	1 (bases 1 to 722)		
AUTHORS	Vettore,A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.		
TITLE	The libraries that made SUCST		
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)		
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bccccenter.fcav.unesp.br Plate: 004 row: G column: 05 Seq primer: T7 Promoter Primer. Location/Qualifiers 1. .722 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCCCT1004G05" /lab_host="DH10B" /clone_lib="RT1" /note="Organ: Root tips (0.3cm-long) from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Root tips (0.3cm-long) from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucst.lad.ic.unicamp.br/public"		
FEATURES	source		
source	1. .722 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCCCT1004G05" /lab_host="DH10B" /clone_lib="RT1" /note="Organ: Root tips (0.3cm-long) from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Root tips (0.3cm-long) from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucst.lad.ic.unicamp.br/public"		
ORIGIN			
Query Match	41.6%;	Score 559;	DB 6; Length 722;
Best Local Similarity	90.9%;	Pred. No. 1.3e-120;	
Matches	639; Conservative	0; Mismatches 60;	Indels 4; Gaps 4;

Qy	37	CCCTTCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGC	96
Db	16	CCTCTCGCCTCACCATGTCGACCT-CGCGCGCGGAGCACCAAGTCGCGGCCACCGC	74
Qy	97	GCTTCGCGCAGCAGCTGGCGCGCTCATCGACGCTCGGCCTCTTCTACAGCCGCTC	156
Db	75	GCTTCGCGCAAGAAGTGGTCCGCTCATCGACGCTCTGGCCTCTTCTACAGCCGCTC	134
Qy	157	CAGCGCGGCGACGTCGGGAGCACAGGTGCGCTTCTATGAGGGTCTCCGCGCCACGCC	216
Db	135	CAGTCTGGGAGACGCGGGGAGCACAGCTCGCCTTCTACGAGGGTCTCCACCCACGCC	194
Qy	217	CGCTCCCGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCAGCGACGCACTCCTC	276
Db	195	CGCGTCCGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCAGCGACGCACTCCTC	254
Qy	277	CCACGAGCGGAGCGCGGGAGCGCATCTCTACCTCTGCTCTCGACGACCTCTCGCG	336
Db	255	CCACGAGCGCGCGGAGCGGAGCGCATCCGCACTCTGTCCTCGACGACCTCTCTCGCG	314
Qy	337	GGGTTTCAGCGCGCTCGTCGACACATCAAGATCGCGCCATCAGTGGCCACCGAGT	396
Db	315	GGCTGGAGCGCGCTTCGCTGCCGACATCAAGATCGCGCGCATCACTGCGCCCGAGC	374
Qy	397	TGCGCGGAGCCCTACATCGCCAAAGTGCTTCGCGCAAGGACCGCGGACCGACGAGGTTCTG	456

[illegible]

RESULT 12	CA202412	CA202412	671 bp	linear	EST 25-SEP-2003
CA202412					
LOCUS		CA202412	671 bp	mRNA	
DEFINITION		SCRLFLU1009H02.g	FL1	Saccharum officinarum	cdna clone SCRLFLU1009H02

5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CA202412  
CA202412.1 GI:35237651  
EST.  
Saccharum officinarum  
Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 671)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenhariaia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
<http://www.bccccenter.fcav.unesp.br>  
Plate: 009 row: H column: 02  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1. .671  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCRLFL1009H02"  
/lab\_host="DH10B"  
/clone\_lib="FL1"  
/note="Organ: Inflorescence at beginning of development  
(1cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;  
An unidirectional cDNA library generated from  
[Inflorescence at beginning of development (1cm-long)].  
cDNA was prepared from polyA+ mRNA using SuperScript  
Plasmid System kit (Invitrogen). The double-strand cDNAs  
were fractionated in a sepharose CL-2B 40cm-column and  
fragments sizing between 0.8 and 1.5 Kb were  
directionally cloned into the vector. Details  
of RNA and library construction can be obtained at  
<http://sucest.lad.ic.unicamp.br/public>"

FEATURES  
source

## ORIGIN

[illegible]

Db	302	TCGGCGTCCGGTCTCGTCCGAGGCGCGTGTGCGGACTGAGCGCCGGAGGTG	361
QY	532	AAGCCATGACACCGCGCGTCCGCGGTCTCCGCGCTACGTGATCCCGTTGCC	591
Db	362	AAGCCCTGACACCGCGCGTCCGCGGTCTCCGCGCTACGTGATCCCGTTGCC	421
QY	592	GACGAGGAGTGAAGTGTGCGTCCGCGCGCGGTGTACGGAGAAAAGTGGAGTCTTG	651
Db	422	GACGAGGAGTGAAGTGTGCGTCCGCGCGCGGTGTACGGCGGCGAAGGGGAGTCTTG	481
QY	652	TCACAGCTCGGAGCTCAAGCGTGTTCGAGGACAGACTCTGTTCACATCTTACTCG	711
Db	482	TCACAGCTCGGAGCTTAAGCGTGTTCGAGGACAGACTCTGTTCACATCTTACTCG	541
QY	712	GCCTCGATTTCTTGGGCTATGATGCTGCTGAGTGCAGCAGCGGAGATGGGGTGGG	771
Db	542	GCCTGGAATCTTTTGGGCTATGATGCTGAGTGCAGCAGCGGAGTGAAGTGGG	601
QY	772	GTGACGCTGAAGCTGTGGAATTTTCCCATGTGCGCGAGGATGGGTGATTGACCAC	831
Db	602	TTGAGGTGAA-CTTGTTGACTTTTCCATTTTGTCAAAGTTATAGGGGGAATGAACCA	660
QY	832	AACTTCTCTGGG	842
Db	661	AAATTTCTGGG	671
RESULT 13			
CC724950/c			
LOCUS	CC724950	722 bp	DNA linear GSS 23-JUN-2003
DEFINITION	OGLAB23TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0307C22, genomic survey sequence.		
ACCESSION	CC724950		
VERSION	1	GI:32143883	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 722) Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N. Consortium for Maize Genomics Unpublished (2002) Other GSSs: OGLAB23TV Contact: Cathy Whitelaw TIGR		
REFERENCE	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends.		
TITLE	Unpublished (2002)		
JOURNAL	Other GSSs: OGLAB23TV		
COMMENT	Contact: Cathy Whitelaw		
FEATURES	source 1. .722 /organism="Zea mays" /mol_type="Genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBma0307C22" /notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"		
ORIGIN	Query Match 40.2%; Score 540.8; DB 9; Length 722; Best Local Similarity 89.8%; Pred. No. 2.4e-116; Matches 638; Conservative 0; Mismatches 17; Indels 57; Gaps 3; QY 246 CTTCCCGCGTTCACGACGCGACTCTCTCCACCGAGCGCGACCGCGGAGCGCA 305 		
Db	722	CTTCCCGCGTTCACGACGCGACTCTCTCCACCGAGCGCGACCGCGGAGCGCA	663
QY	306	TCCTCACCTCGTCTCGAGGACCTCTCGCGGGTTTCAGGCGCCCTGGTGCAGACAT	365
Db	662	TCGACACCTCGTCTCGAGGACCTCTCGCGGGTTTCAGGCGCCCTGGTGCAGACAT	603
QY	366	CAAAGATCGGCGCATCAGTGGCCACCGAGTTTCGCGGAGCCCTATATCGCCAAAGTGCCT	425
Db	602	CAAAGATCGGCGCATCAGT-----	583
QY	426	CGCCAAGACCGCGGGACCAAGAGCTTCTGCTCGGATTCGCGGTCCGAGT	485
Db	582	-----GACCAGAGGTTCTGCTGGATTCGCGGTCTCCGCGTCCGAGT	538
QY	486	CGTCCGCCCCGAGGCGCGGTGTGCGGACGAGAGCGCCGCGAGGTGAAGCCCATGACAC	545
Db	537	CGTCCGCCCCGAGGCGCGGTGTGCGGACGAGAGCGCCGCGAGGTGAAGGCTATGACAT	478
QY	546	CGCGCGCGTCCGCGCGGTCTCCGCGGTACGTGTCATCCG--TTGCCGACGAGGGATGG	604
Db	477	TGTCGCGGTCCGCGCGGTCTCCGCGGTACGTGTCATCCGTTGCCGAGAGGGATGG	418
QY	605	ACTGTGCGTCTCGCCGCGCGGTGTACGAGGAAAAAGTGGAGTCTTTGTACAGCTGCGCG	664
Db	417	ACTGCGCGTCCGCGCGCGGTGTACGAGGAAAAAGTGGAGTCTTTGTACAGCTGCGCG	358
QY	665	AGTCAAGCGGTGTTGAGGAGCAGACTCTGTTCCACTTTCTACTCGGCGTGCATTTCTTC	724
Db	357	AGTCAAGCGGTGTTGAGGAGGCGAGACTCTGTTCCACTTTCTACTCGGCGTGCATTTCTTC	298
QY	725	TGGGCTATGATCTCTCGAGTCGAGCGGAGATGGGGTGGGTGACGCGTGAAGC	784
Db	297	TGGGCTATGATCTCTCGAGTCGAGCGGAGGTGGGGTGGGTGAACAGTGAAGC	238
QY	785	TGGTGGACTTTGCCCATGTGCGCGAGGGTGAATGGGCTGATTGACCACAACTCTCTGGGCG	844
Db	237	TGGTGGACTTTGCCCATGTGCGCGAGGGTGAATGGGCTGATTGACCACAACTCTCTGGGCG	178
QY	845	GGCTCTGCTCGTGTATCAAGTTCGTTTCTGACATTTGTTCCGGAGACTCTCTATACGCGC	904
Db	177	GGCTCTGCTAGTGTATCAAGTTCGTTTCTGACATTTGTTCCAGAGACTCTCTCAGACGCGC	118
QY	905	CTTTGGGTCCTTTCTTAAGAGAGGATCTCTGGCA--TTTCGATTTGATAACAAG 955	
Db	117	CTTTGGGTCCTTTCTTAAGAGAGGATCTCTGGCA--TTTCGATTTGATAACAAG 66	
RESULT 14			
BZ774817			
LOCUS	BZ774817	593 bp	DNA linear GSS 14-MAR-2003
DEFINITION	ii51a11.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone ii51a11, genomic survey sequence.		
ACCESSION	BZ774817		
VERSION	BZ774817.1	GI:28952260	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 593) Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N., Katzenburger, P., King, L., Miller, B., Muller, S., Nascento, L., Zukavern, T., McCombie, W.R. and Martienssen, R.A. Genomic shotgun sequences from Zea mays (methyl1-filtered) Unpublished (2002) Contact: W. Richard McCombie Lita Annenberg Hazen Genome Sequencing Center Cold Spring Harbor Laboratory PO Box 100, Cold Spring Harbor, NY 11724, USA Tel: 516 367 8884 Fax: 516 367 8874 Email: mccombie@cshl.org		
REFERENCE	1 (bases 1 to 593)		
AUTHORS	Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N., Katzenburger, P., King, L., Miller, B., Muller, S., Nascento, L., Zukavern, T., McCombie, W.R. and Martienssen, R.A.		
TITLE	Genomic shotgun sequences from Zea mays (methyl1-filtered)		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: W. Richard McCombie		



Plate: ii51 row: a column: 11  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 593.  
Location/Qualifiers  
1. 593

FEATURES  
source

/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="i151a11"  
/lab\_host="DH5a"  
/clone\_lib="WGS-ZmaysF (DH5a methyl filtered)"  
/notes="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector (.x/y  
reads in M13mp19, b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match 39.9%; Score 535.6; DB 8; Length 593;  
Best Local Similarity 97.9%; Pred. No. 3.9e-115;  
Matches 564; Conservative 0; Mismatches 9; Indels 3; Gaps 2;  
QY 42 TCCCATATACATGTCGACCTCCACCCGCGGAGACCAAGTCGCGGCACCGCGCTC 101  
DB 20 TCCCATATACATGTCGACCTCCACCCGCGGAGACCAAGTCGCGGCCTC 79  
QY 102 CGCAGCAAGCTGGCGCGCTCATGACGCTCGGCGCTTCTTACAAAGCGTCCAGGC 161  
DB 80 CGCAGCAAGCTGGCGCGCTCATGACGCTCGGCGCTTCTTACAAAGCGTCCAGGC 139  
QY 162 CGCGACCGTGGGAGACGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGCGCT 221  
DB 140 CGCGACCGTGGGAGACGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGCGCT 199  
QY 222 CCGGCGCGCATCGAGACCTTTTCCCGGTTCCAGGACCGGACGACTCTCTCCCGAC 281  
DB 200 CCGGCGCGCATCGAGACCTTTTCCCGGTTCCAGGACCGGACGACTCTCTCCCGAC 259  
QY 282 CGAGGCGCAGCGCGGAGCGGATCTTCACTCGCTCTCGACGACTCTCTCGGGGTT 341  
DB 260 CGAGGCGCAGCGCGGAGCGGATCTTCACTCGCTCTCGACGACTCTCTCGGGGTT 319  
QY 342 TCAGGCGCCTTGGTTCGACATCAAGATCGGCGCCATCACTGCGGACCGGACTTCGCC 401  
DB 320 TGAGGCGCCTTGGTTCGACATCAAGATCGGCGCCATCACTGCGGACCGGACTTCGCC 379  
QY 402 GGAGCCCTACATCGCCAAAGTCTTCCGCAAGGACCGCGGACCAAGAGCGTTCTGTCGG 461  
DB 380 GGAGCCCTACATCGCCAAAGTCTTCCGCAAGGACCGCGGACCAAGAGCGTTCTGTCGG 439  
QY 462 ATTCCGGCTTCCGGCTTCGAGTCTCGGCGCGCGGCGCGCTGTGGCGGAGCGAGCG 521  
DB 440 ATTCCGGCTTCT--TGCCTCGAGTCTGTGGCGCGCGGCGCGCTGTGGCGGAGCG 497  
QY 522 CCGGAGGTGAAGCCATGACACCGCGCGCTCGCGCGCTGTCTCGGCGCTACGTGTC 581  
DB 498 CCGGAGGTGAAGCTATGACACCGGTGCGGCTCCGCGGCTGTCTCGGCGCTACGTGTC 557  
QY 582 ATCCG-TTGGCGAGGAGGTGACTGTGGCTCG 616  
DB 558 ATCCGCTTGGCGAGGAGGTGACTGTGGCTCG 593

RESULT 15  
CAL130686 744 bp mRNA linear EST 24-SEP-2003  
LOCUS SCCCRT1004G06.g Rt1 Saccharum officinarum cDNA clone SCCCRT1004G06  
DEFINITION 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CAL130686  
CAL130686.1 GI:35013916  
EST.

## SACCHARUM OFFICINARUM

Saccharum officinarum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.

## REFERENCE

1. (bases 1 to 744)

## AUTHORS

Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.

## TITLE

The libraries that made SUCESI

## JOURNAL

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

## COMMENT

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bccccenter.fcav.unesp.br

Plate: 004 row: G column: 06

Seq primer: T7 Promoter Primer.

FEATURES  
source

Location/Qualifiers  
1. 744

/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCCRT1004G06"

/lab\_host="DH10B"

/clone\_lib="RT1"

/notes="Organ: Root tips (0.3cm-long) from adult plants;  
vector: pSPori; Site\_1: SalI; Site\_2: NotI; An

unidirectional cDNA library generated from [Root tips

(0.3cm-long) from adult plants]. cDNA was prepared from

polyA+ mRNA using SuperScript plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucest.lad.ic.unicamp.br/public"

## ORIGIN

Query Match 39.1%; Score 525.6; DB 6; Length 744;  
Best Local Similarity 88.6%; Pred. No. 9e-113;  
Matches 592; Conservative 0; Mismatches 74; Indels 2; Gaps 2;  
QY 37 CCCTTCCCATACCATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGGCACCGC 96  
DB 16 CCTCTCGCCTTACCATGTCGACCT-CCGCCCGCGGAGCACCAAGTCGCGGCACCGC 74  
QY 97 GCCTCCGCGCAGCAAGTGGCGCGCTCATCGAGGCTCCGCGCTCTTCTACAGCGCTC 156  
DB 75 GCCTCCGCGCAGCAAGTGGCTCGCTCATCGAGGCTCTGCGCTCTTCTACAGCGCTC 134  
QY 157 CAGCGCGCGGACCGTGGGAGCAGAGGTGCGCTTCTATGAGCGGTTCTCCGCCAGCGC 216  
DB 135 CAGGTCCGGGACCGCGGGGAGCAGAGTCTGCGCTTCTACGAGGCTTCTCACCACCGC 194  
QY 217 GCGTCCGCGCGCGATCCGAGACACTTCTCCCGGTTCCACGCGACGCGACTCTC 276  
DB 195 GCGTCCGCGCGCGATCCGAGACACTTCTTCCCGGTTTCCGCGACGCGACTCTC 254  
QY 277 CCACCGAGGCGCAGCCCGGAGGAGCGCATCTTCTTCCCGGTTTCCGCGACGCGACTCTC 336  
DB 255 CCACCGAGGCGGCGGAGGAGGAGCGCATCCGACCTCTCTCGAGGACTCTCTCGG 314  
QY 337 GGGTTTTCAGCGCGCTTCGTCGAGACATCAAGATCGCGCCATACGTCGGCCACCGAGT 396  
DB 315 GGGTGGAGGCGCGCTTCGTCGCGGACATCAAGATCGCGCCATACGTCGGCCCGGAGC 374

Qy	397	TCGCCGAGCCCTACATCGCAAGTGCCTCGCCAGGACCGCGGGACCAAGCGTTCTG	456
Db	375	TCGCCGAGCCCTACATCGCCAAAGTGCCTCGCCAGGACCGCGGGACCAAGCGTTCTG	434
Qy	457	CTCGGATTCCGGGTCTCCGGCGTCCGAGTGTGCGGCCCGAGGGCCCGTGTGGCGGACG	516
Db	435	CTCGGATTCCGGGTCTCCGGCGTCCGGGTCTCGGTCCCGAGGGCCCGTGGGGGGACT	494
Qy	517	GAGCGCCCGAGGTGAAGGCCATGGACACGCGCGCGTCCGCGCGTGTCTCCGGCGCTAC	576
Db	495	GAAAGCGCCCGAGGTGAAGGCCCTTGGACACCGCGCGGTCCCGCGGGGTTCGGCGCTAC	554
Qy	577	GTGTCATCCGTTGCCGACGAGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGAGGA	636
Db	555	GTTTTATTCTGTGCCGACGAGGGATGGACTGTCTCCCTCGCGCGCGGT-TACGGGGC	613
Qy	637	AAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGTGTTCAGGAGCAGACTCTG	696
Db	614	AAAGGGGAAACTTGGTAAACATGGGCAAGCTTAAAGGGGGGTAGAGGGGAAATTTGG	673
Qy	697	TTCCACTT	704
Db	674	TTCCTTT	681

Search completed: June 17, 2005, 01:52:45  
Job time : 4778 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 22:35:02 ; Search time 254 Seconds  
(without alignments)  
8658.094 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....atacaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69.6	5.2	1926	3	US-09-249-585A-4
2	69.6	5.2	1931	2	US-09-130-114-2
3	58.6	4.4	38506	3	US-09-320-878-19
4	58.6	4.4	38506	4	US-09-141-908-1
5	58.6	4.4	38506	4	US-09-657-440-19
6	58.6	4.4	4411529	3	US-09-103-840A-1
7	58.4	4.3	18235	4	US-09-949-016-13686
8	58.4	4.3	22547	4	US-09-949-016-13679
9	58	4.3	1082	4	US-09-881-165-4
10	58	4.3	1680	4	US-09-902-540-7284
11	58	4.3	5054	4	US-09-902-540-687
12	57.6	4.3	4530	4	US-09-799-451-881
13	57.6	4.3	23673	3	US-09-773-816-1
14	57	4.2	11220	3	US-09-105-537-32
15	57	4.2	36778	3	US-09-105-537-5
16	56.8	4.2	8563	4	US-09-902-540-3318
17	56.8	4.2	15351	4	US-09-902-540-1154
18	56.4	4.2	2595	4	US-09-902-540-3202
19	56.4	4.2	19455	4	US-09-902-540-1147
20	55.2	4.1	15872	3	US-09-105-537-1
21	55.2	4.1	15872	4	US-09-091-609-1
22	55.2	4.1	15872	4	US-09-091-609-3
23	55.2	4.1	4403765	3	US-09-103-840A-2
24	55	4.1	1294	4	US-10-151-832-4
25	54.8	4.1	1704	1	US-08-528-199-2
26	54.8	4.1	1704	1	US-08-528-199-5
27	54.8	4.1	4403765	3	US-09-103-840A-2

C	28	54.4	4.0	985	3	US-09-056-556-182	Sequence 182, App
C	29	54.4	4.0	985	3	US-09-072-596-177	Sequence 177, App
C	30	54.4	4.0	985	4	US-09-072-967-182	Sequence 182, App
C	31	54.4	4.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C	32	54.2	4.0	1086	4	US-09-902-540-7871	Sequence 7871, Ap
C	33	54.2	4.0	6196	4	US-09-902-540-788	Sequence 788, App
C	34	53.8	4.0	1333	3	US-09-372-422A-9	Sequence 9, Appli
C	35	53.4	4.0	921	4	US-09-252-991A-6598	Sequence 6598, Ap
C	36	53.4	4.0	1143	4	US-09-902-540-3850	Sequence 3850, Ap
C	37	53.4	4.0	3084	4	US-09-252-991A-6639	Sequence 6639, Ap
C	38	53.4	4.0	3273	4	US-09-252-991A-6578	Sequence 6578, Ap
C	39	53.4	4.0	9556	4	US-09-902-540-982	Sequence 982, App
C	40	53.4	4.0	21964	4	US-09-902-540-1190	Sequence 1190, Ap
C	41	53.2	4.0	1419	4	US-09-902-540-130	Sequence 130, App
C	42	53.2	4.0	6854	3	US-09-194-905-7	Sequence 7, Appli
C	43	53.2	4.0	6854	4	US-09-922-683-7	Sequence 7, Appli
C	44	53.2	4.0	152331	3	US-09-128-155-16	Sequence 16, Appli
C	45	52.6	3.9	810	4	US-09-902-540-3387	Sequence 3387, Ap

## ALIGNMENTS

## RESULT 1

US-09-249-585A-4  
; Sequence 4, Application US/09249585A  
; Patent No. 6417002  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert  
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES  
; FILE REFERENCE: 0867/0D905  
; CURRENT APPLICATION NUMBER: US/09/249,585A  
; CURRENT FILING DATE: 1999-02-11  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 1926  
; TYPE: DNA  
; ORGANISM: Epstein Barr Virus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(1926)  
; OTHER INFORMATION: template strand of EBNA-1 DNA  
US-09-249-585A-4

Query Match		5.2%	Score 69.6;	DB 3;	Length 1926;
Best Local Similarity		45.8%	Pred. No. 7.8e-07;		
Matches 246;		Conservative	0;	Mismatches 294;	Indels 0; Gaps 0;
QY	39	CCTTCCCATACATGTCGACCTCCACCGCGGAGACCAAGTCGCGCGCCACCGCGC	98		
DB	274	CCTCGTCCTCGCGCTCCCGCTCTCGCTCTCCCGCTCTCGCTCTCTCTCTCTCTCTCT	333		
QY	99	CTCCGCCAGACGCTGGCGCGCTCATGACGGCTCTGGGCTCTTTTACAGCGCTTCCA	158		
DB	334	CCTCTCCCGCTCTCCCGCTCTCCCGCTCTCCCGCTCTCTCTCTCTCTCTCTCTCTCT	393		
QY	159	GGCGCGGACGCTGGGAGACAGAGTCTGCTTCTATAGGCGTCTTCGCGCCACCGCGC	218		
DB	394	CGTCTCTCTCCCGCTCTCCCGCTCTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	453		
QY	219	CGTCCCGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCAGCGACGCGACTCTCTCCC	278		
DB	454	CCTCCCGCTCTCTCCCGCT	513		
QY	279	CACCGAGGCGAGCCCGGGAGCCGATCCTTACCTCTCTCTCGAGACCTCTCTCGGGG	338		
DB	514	CCTCTCCCGCT	573		
QY	339	GTTTACGGCGCTCTGCTCGCAGACATCAGATCGGCGCATCAGTGGCCACCGAGTTC	398		
DB	574	CCCGGCTCTCCCGCT	633		

[illegible]

## RESULT 2

```

US-09-130-114-2
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Basam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

```

### RESULT 3

```

US-09-320-878-19
; Sequence 19, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-878-19

```

```
Db 22631 CGCGCCGACGGCCTCG 22647
|||||
RESULT 4
US-09-141-908-1
; Sequence 1, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141.908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-141-908-1
Query Match 4.4%; Score 58.6; DB 4; Length 38506;
Best Local Similarity 44.9%; Pred. No. 0.00087;
Matches 223; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

Qy 66 CCGCGCGGACCAAGTCGCGGCCACCGCGCCTCGCGCAGCAAGTCGGCCCGCTCAT 125
|||||
Db 22151 CCGCGCGGACCGTCCTGCTGGCTGCGCCCGCGCCCGCGTGGCGGAGCATGTCG 22210
|||||
Qy 126 CGAGCGCTCGGCTCTTCTACAAGCGCTTCCAGGCGCGCGACCGTGGGAGCACGAGGT 185
|||||
Db 22211 CGAGGCGCTGACCGGTCGCTCGGCTGATGAGGCTGCTGGCCGACGAGCGTTTAC 22270
|||||
Qy 186 CGCCTTCTATAGGCGTTCTCGCCCAACCGCGCTCCCGCGCGCATTCGAGACACTT 245
|||||
Db 22271 CGATGGCGCTGCTGCTGACCGCGCGCTGCGCGCTCGCGCGCGAGAGGCC 22330
|||||
Qy 246 CTTCCCGCGTTCCAGGACGCGACTCTCTCCCAACGAGGCGCGCGCGCGCGCGCGCA 305
|||||
Db 22331 CGGCTCACGGGACAGCGCGCTGCGCGCTGCGCGCTGCGCGCGAGAGGCC 22390
|||||
Qy 306 TCCTACCTCGTCTCGACGACCTCTCGCGCGGTTTCAGGCGCTTCGCGCAGACAT 365
|||||
Db 22391 GGGCGGTTGCTGCTGCTGACCTCGCGGGGAAGCCCGAGCGCGCGGAGCGCCACCG 22450
|||||
Qy 366 CAAGATCGGGCCCATCAGTGGCCACCGAGTTCCCGGAGCCCTTACATCGCCAAAGTGCCT 425
|||||
Db 22451 CGGGGACGGCTGACGACCGGGGAGCGCACCGTCCGGGCTTCAGGCGCTTCGCGCAGACAT 485
|||||
Qy 426 CGCAGAGGACCGGGGACCGAGCGTTCTGCTCGGATTCGGGCTTCGGCGCTCCCGGCTCCGAGT 485
|||||
Db 22511 CGGCGGCTTCGCTGCTGCTGACCTCGCGGGGAAGCCCGAGCGCGCGCTTCGCGGACGG 22570
|||||
Qy 486 CGTGGCGCGCGCGCGCTGTCGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
|||||
Db 22571 GGGCGCTCTCTGTAACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
|||||
Qy 546 CGCGCGGCTCCGCGCG 562
|||||
Db 22631 CGCGCGGACGGCCTCG 22647
|||||

RESULT 6
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
Query Match 4.4%; Score 58.6; DB 4; Length 38506;
Best Local Similarity 44.9%; Pred. No. 0.00087;
Matches 223; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

Qy 66 CCGCGCGGACCAAGTCGCGGCCACCGCGCCTCGCGCAGCAAGTCGGCCCGCTCAT 125
|||||
Db 22151 CCGCGCGGACCGTCCTGCTGGCTGCGCCCGCGCCCGCGTGGCGGAGCATGTCG 22210
|||||
Qy 126 CGAGCGCTCGGCTCTTCTACAAGCGCTTCCAGGCGCGCGACCGTGGGAGCACGAGGT 185
|||||
Db 22211 CGAGGCGCTGACCGGTCGCTCGGCTGATGAGGCTGCTGGCCGACGAGCGTTTAC 22270
|||||
Qy 186 CGCCTTCTATAGGCGTTCTCGCCCAACCGCGCTCCCGCGCGCATTCGAGACACTT 245
|||||
Db 22271 CGATGGCGCTGCTGCTGACCGCGGAGCGGTCGCGCGCTTCGCGGAGCGCCT 22330
|||||
Qy 246 CTTCCCGCGTTCCAGGACGCGACTCTCTCCCAACGAGGCGCGCGCGCGCGCGCGCA 305
|||||
Db 22331 CGGCTCACGGGACAGCGCGCTGCGCGCTGCGCGCTGCGCGCGAGAGGCC 22390
|||||
Qy 306 TCCTACCTCGTCTCGACGACTCTCTCGGGGTTTCAGGCGCTTCGCGCAGACAT 365
|||||
Db 22391 GGGCGGTTGCTGCTGCTGACCTCGCGGGGAAGCCCGAGCGCGCGGAGCGCCACCG 22450
|||||
Qy 366 CAAGATCGGGCCCATCAGTGGCCACCGAGTTCCCGGAGCCCTTACATCGCCAAAGTGCCT 425
|||||
Db 22451 CGGGGACGGCTGACGACCGGGGAGCGCACCGTCCGGGCGGACCTCTGAGAGCGCGCCT 22510
|||||
Qy 426 CGCAGAGGACCGGGGACCGAGCGTTCTGCTCGGATTCGCGCTTCGCGGCTCCGAGT 485
|||||
Db 22511 CGGCGGCTTCGCTGCTGCTGACCTCGCGGGGAAGCCCGAGCGCGCGCTTCGCGGACGG 22570
|||||
Qy 486 CGTGGCGCGCGCGCTGTCGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
|||||
Db 22571 GGGCGCTCTCTGTAACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
|||||
Qy 546 CGCGCGGCTCCGCGCG 562
|||||
Db 22631 CGCGCGGACGGCCTCG 22647
|||||
```

Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 4.4%; Score 58.6; DB 3; Length 4411529;  
Best Local Similarity 44.4%; Pred. No. 0.0042;  
Matches 235; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 82 GTGCGCGGCACCGCGCTCCGCGCAGCAAGCTGGGCGCGCTCATCGACGGCTCCGCGCTC 141  
Db 3934336 GTGCGCGGCACCGCGCTCCGCGCAGCAAGCTGGGCGCGCTCATCGACGGCTCCGCGCTC 3934277

QY 142 TTCTACAAGCGCTCCAGGCGCGCGACCGCTGGGAGCAGAGTGCGCTTCTATGAGGCG 201  
Db 3934276 GTGCGCGGCCTTGGCTGTGAGGCGGACCGCGCGCTTGGCGCGTTCGCGCGGCGCAC 3934217

QY 202 TTCTCGGCGCAGCGCGCTCCGCGCGCGATCCGAGACACTTCTCCCGGTTCCAC 261  
Db 3934216 GGCGCGCGCTCAACCGCGCGCGCTTGGCGCGCTTGGCGCGAGACTGCGCGCACCGTTGCGCC 3934157

QY 262 GGCACGCGACTCTCCCGCAGGCGCGAGCGCGGAGCGCGATCTCACCTCGTCTC 321  
Db 3934156 GTCCCGCGCTGCGCGCGCTTGGCGCGCGCGCGCGCGCTTGTATGCGCGCGCGC 3934097

QY 322 GACGACCTCTCTCGCGGGTTTACGGCGCGCTTGGCGCGCGATCAAGATCGCGCGCATC 381  
Db 3934096 GGCGCTGCGCGCGCGCGCTTGGCGCGCTTGGCGCGCGTCAAGCGCGAGGCGCGAG 3934037

QY 382 ACGTGCGCGCGAGTTGCGCGGAGCGCTACATCGCGCGCGCTTGGCGCGCGCGCGG 441  
Db 3934036 GCGGAGACCGCTGGCGCGCGCTTGGCGCGCGCGCGCGCGCTTGGCGCGCGCGCGCG 3933977

QY 442 ACCACGAGCGTTCTGCTCGGATTCGCGCTTCCGCGGCTCGGAGTCTCGCGCGCGCGAGGCG 501  
Db 3933976 GGCGTTACCG 3933917

QY 502 GCGGTGCGCGGAGCG 561  
Db 3933916 CCGCGCGGTGCG 3933857

QY 562 GTGCTCGCGCGCTACGTGTCTATCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 610  
Db 3933856 GCGCGCGGTGCG 3933808

RESULT 7  
US-09-949-016-13686/c  
; Sequence 13686, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13686  
; LENGTH: 18235  
; TYPE: DNA  
; ORGANISM: Human  
; OTHER INFORMATION: H37Rv  
US-09-949-016-13686

Query Match 4.3%; Score 58.4; DB 4; Length 18235;  
Best Local Similarity 46.2%; Pred. No. 0.00076;  
Matches 194; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 186 CGCCTTCTATGAGGCTTCTCGCGCCAGCGCGCGCTCCGCGCGCGCTCCGCGCGCGCTCCGCT 245  
Db 18021 CGGCTTGGCGGACCCACCTCGGTGCGCTCCGCGCGCGCGCTCCGCTCGGCTGGGCT 17962

QY 246 TTCTCCCGGTTCCAGCGACGCGACTCTCTCCCGCGCGCGCTCCGCGCGCGCTCCGCGCGCGCT 305  
Db 17961 CGCGCGCGCTTCCAGCGCGCGCGCTCTCCCGCGCGCGCTCCGCGCGCGCTCCGCGCGCGCT 17902

QY 306 TCCTACCTCTCTCGAGGACCTCTCGCGGGTTTTCAGGGCGCTTTCGCGCGCGCTTTCGCGAGCAT 365  
Db 17901 GCGTCCCGCGCTTCCCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTT 17842

QY 366 CAAGATCGCGCGCATCAGCTGGCCACCGAGTTTCGCGCGCGCGCTTACATCGCCCAAGTGCCT 425  
Db 17841 CACCGGGGCTCCCTCTCGCGCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTT 17782

QY 426 GCGCAAGACCGCGCGGACCGAGGCTTCTCGTCTCGGATTCGCGCTTCCGCGCTTCCGCGCTTCCGAGT 485  
Db 17781 TGCGCGCGCGCGCGCGCGCGCTTCCGCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCG 17722

QY 486 GTGCGCGCGCGCGCGCGCTTTCG 545  
Db 17721 GCG 17662

QY 546 GCGCGCGCTCGCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCGCGCGCG 605  
Db 17661 CTCTCCCG 17602

RESULT 8  
US-09-949-016-13679/c  
; Sequence 13679, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13679  
; LENGTH: 22547  
; TYPE: DNA  
; ORGANISM: Human  
; OTHER INFORMATION: H37Rv  
US-09-949-016-13679

Query Match 4.3%; Score 58.4; DB 4; Length 22547;  
Best Local Similarity 46.2%; Pred. No. 0.00081;

Matches	194;	Conservative	0;	Mismatches	226;	Indels	0;	Gaps	0;
Qy	186	CGCCTTCTATGAGCGGTTCTCCGCCACGCGCGCTCCCGGCCGATCCGAGACACCTT	245						
Db	2686	CGGCTTGGCGGACCCCACTTCGCTGCGGTCCCGCCGCGGCGCTTCGCTTGGGCTGCGGCT	2627						
Qy	246	CTTCCCCCGGTTCCACGGCACGCGACTCTCTCCACACGAGCGCAGCCCGGGAGCGCGCA	305						
Db	2626	CCCCCGCGCTCTCACGCGGGGGCCCGCTCCCGCGCCCGCCCGCGCCCGCGCCA	2567						
Qy	306	TCCTCACCTCGTCTCGACGACCTCTCTCGGGGGTTTTAGGCGCCCTGCGTCGAGACAT	365						
Db	2566	GGCTGCCCGCTTCCCGCTCTCGCCCTTGGCGGGAGCTCGGCACCGCGGCTGGCGCACCCC	2507						
Qy	366	CAAGATCGGGGCCATCAGTGGCCACCGAGTTTCGCGGAGCGCTTACATCGCAAGTCCT	425						
Db	2506	CACCGGGGCTCCCTCTCGCCCCGCA CGCGCGGTCTCTGGGCGAGTTTAAACCTCTCCC	2447						
Qy	426	CGCCAAGAACCGCGGAGCCACGAGCGTTTGTCTCGGATTCGCGGTCTCCGGCGTCCGAGT	485						
Db	2446	TGCCCGCGCGCCGGGAGGCCAGGGCCGCGGGCGGGGGCTGCGCCACTAGTCCCCCGA	2387						
Qy	486	CGTCGGGCCCGAGGGCCGTGTGTGGCGGA CGGACGCCCCCGAGGTGAAGGCCATGGACAC	545						
Db	2386	GGGCGGGCGACCGGAGGCCCGCGCCCGGACGCGGAGGGGACCGAAACGAAGAGCC	2327						
Qy	546	CGCCGGCGGTCCGCGGTGTCTCGGCGGTACGTGTCTATCTCGCTTGCACGAGGGGATGGA	605						
Db	2326	CCTCCCCCAACCGCGGAGCGCCCACTGCCCCCGCCCTGCTCGGTGTTCGGCGCGAGGAA	2267						

```

RESULT 9
US-09-881-165-4
; Sequence 4, Application US/09881165
; Patent No. 6632930
; GENERAL INFORMATION:
; APPLICANT: HOOD, ELIZABETH
; APPLICANT: HOWARD, JOHN
; APPLICANT: BAILEY, MICHELE
; APPLICANT: GASTEL, FRANS VAN
; APPLICANT: WANG, HUAMING
; APPLICANT: WARD, MICHAEL
; APPLICANT: WOODARD, SUSAN
; TITLE OF INVENTION: METHOD OF INCREASING RECOVERY OF HETEROLOGOUS ACTIVE
; FILE OF INVENTION: ENZYMES PRODUCED IN PLANTS
; FILE REFERENCE: 10032R
; CURRENT APPLICATION NUMBER: US/09/881,165
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/211,732
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1082
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA encoding
; OTHER INFORMATION: Organophosphate Hydrolase
US-09-881-165-4

```

	Query Match	4.3%	Score 58;	DB 4;	Length 1082;
	Best Local Similarity	46.2%;	Pred. No. 0.00037;		
	Matches 230;	Conservative	0;	Mismatches 265;	Indels 3; Gaps 1;
QY	11	AGTCGGTCACCGCTCGCGGCCCATGATCCCTCTCCCATACCATGTCGCACTCCACCCGC	70		
Db	121	AGCGCGGTTACCTCCACCCACGAGACATCTGGGCTCTCCGCGGCTTCTCCGCG	180		
QY	71	CGGAGCACAGTGTCCCGGCCACCGCGGCTTCGGCCAGCAAGCTGGCGCGGTCTATCGACG	130		
Db	181	CTGGCCGGAGTTCTTCGGCTCCCGCAAGGGCCCTCGCGAAGAGCGGTGGCGGCTCC	240		

Qy	131	GCTCGCGCTCTTCTATACAGCCGCTCCAGCGCGCGACCGTGGGGAGACACGAGTTCGCGCT	190
Db	241	GCGCGCGCGCGCGCGCGCGTGCACCATCGTGACGTGTCCACCTTCGACATCGGCC	300
Qy	191	TCATAGGCGGTTCCTCGGCCACGCGCGCTCCCGGCCCGCATCCGAGACACCTTCTTTC	250
Db	301	GCGACGTGTCCCTCTCTCGCGGAGTGTCCCGCGCGCGACGTGCACATCGTGGCGGCCA	360
Qy	251	CCCGGTTCACGGCAGCGGACTCTCTCCNCAGAGGGGACGCCGGGAGCCGCATCCTC	310
Db	361	CCGCGCTCTGTGTCGACACCGCGCCTCTCCATGCGCCTCCGCTCCGTGGAGGAGCTCACCC	420
Qy	311	ACCTCGTTCCTCGACGACCTCTCTCGCGGGTTCAGCGCGCCTCGGT---CGCAGACATCA	367
Db	421	AGTTCTTCCTCCCGGAGATCCAGTAGGCATCGAGACACCGGCATCCGCGCGGCATCA	480
Qy	368	AGATCGCGGCATCAGCTGSCACACCGAGTTCGCGGAGCCCTTACATGCGCAAGTGCTCG	427
Db	481	TCAAGTGGCCACCACCGCGGACCCCGCTTCAGGAGCTCGTGCTCAAGGCGCGCG	540
Qy	428	CAAAGGACCGCGGACCAAGAGGTTCGTCTCGGATTCGCGCTCTCGGCGTCCGAGTCG	487
Db	541	CCGCGCTCTCTTCGCCACCGCGGTCCGGTAGCACCCACACACCGCGCCTCCGAGCGCG	600
Qy	488	TCGSCCGCGAGGCGCGC 505	
Db	601	ACGCGGACGACGAGCGCG 618	

RESULT 10  
 US-09-902-540-7284  
 ; Sequence 7284, Application US/09902540  
 ; Patent No. 6833447  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Wiegand, Roger C.  
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
 ; FILE REFERENCE: 38-10(15849)B  
 ; CURRENT APPLICATION NUMBER: US/09/902,540  
 ; CURRENT FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: 60/217,883  
 ; PRIOR FILING DATE: 2000-07-10  
 ; NUMBER OF SEQ ID NOS: 16825  
 ; SEQ ID NO 7284  
 ; LENGTH: 1680  
 ; TYPE: DNA  
 ; ORGANISM: Myxococcus xanthus  
 ; US-09-902-540-7284

Query Match	4.3%;	Score 58;	DB 4;	Length 1680;
Best Local Similarity	48.2%;	Pred. No. 0.00042;		
Matches 163;	Conservative 0;	Mismatches 175;	Indels 0;	Gaps 0;
QY	282	CGAGGCGACGCGGGAGCGCATCTCACTCGTCTCTCGAGACCTCTCGCGGGGTT	341	
DB	1074	CAAGGCCCTTCCCAAGCGCGCGCCCTCATCTTTGCGCGCTGGCGCGCGAGGA	1133	
QY	342	TCAGGCGCCTCGCTCGCAGACATCAAGATCGGGCCCATCAGTGGCCACCGAGTTGCCC	401	
DB	1134	GCAGGGCTGTGSGCTCGAGTACTGTGGCGGAGCACCGCGCTGCCCAACGCGCGGT	1193	
QY	402	GGAGCCTCATATCGCCAAAGTGCTCTGCCAAGGACCGGGGACACACGAGCGTTCTGCTCGG	461	
DB	1194	CGCGCCAAACATCAACATCGACGCGGCCAATCCATGGCGGACCCCGCGACCTCACCGT	1253	
QY	462	ATTCGGCGTCTCGGSGTCCGAGTCGTGCGGCCCGAGGGCGCTGTGGCGGACGCGAGCG	521	
DB	1354	CATTGGCCTGGGCAATCTCAACCTGGACGCCACACTGGTGCCTCTGGCGAAGACGCGGG	1313	
QY	522	CCCGGAGGTGAAGGCCCATGGACACCGCGCGGCTCCGCGCGGTCTCGGGCGCTACGTGTC	581	

Db 1314 CCGGTGTGTAAGCGGACAGCTGTGCGACCGCGCTTCTTCTACCGGTGCGACCAAGTT 1373  
QY 582 ATCCGTTGCGGAGGAGGTGACCTGTCGGCTCGCGG 619  
Db 1374 CAACTTCGCAAGCGGGCATTCGCCCGGTACTTCG 1411

## RESULT 11

US-09-902-540-687/c  
; Sequence 687, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 50/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 687  
; LENGTH: 5054  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-687

Query Match 4.3%; Score 58; DB 4; Length 5054;  
Best Local Similarity 48.2%; Pred. No. 0.00061;  
Matches 163; Conservative 0; Mismatches 175; Indels 0; Gaps 0;  
QY 282 CGAGGCGACCGCGGAGCGCATCTCTCCTCGAGACCTCTCTCGGGGTT 341  
Db 1955 CAAAGGCCCTTCCCAAGGCGCGCGCTCCATCTCTTTCGCGCGTGGCGCGGAGA 1896  
QY 342 TCAGGGCCCTGGGTGCGAGACATCAAGATCGCGCATACGTGSCCACCAGTTGCGC 401  
Db 1895 GCAGGGGCTGTGGGTTCGAGTACTGTGCGGAGACCCCGCGTGCCACCAGCGGGGT 1836  
QY 402 GGAGCCCTACATGCGCAAGTGCTTCGCGAAGGACCGCGGAGACAGAGCGTTCTGCTCGG 461  
Db 1835 CGCGGCAACATCAACATCGAGCGGCGCAACATCCATGSCGACCCGCGACCTCACGT 1776  
QY 462 ATTCCGCTCTCGGGCTCGAGTGTGTCGCGCGCGGCGCGGTGTGCGGACGAGCG 521  
Db 1775 CATTGCCCTGGGCAATTCACCTTGAGCGCCACACTGTGTGCGCCTGGCGAAGACGAGG 1716  
QY 522 CCGGAGGTGAAGGCCATGGACACCGCGCGTTCGCGCGCGTCTCGGCGCTACGTGTC 581  
Db 1715 CCGGTGTGTGAAGCGGACCACTGTGCGACCGCGGCTTCTTCTACCGGTGCGACCAAGTT 1656  
QY 582 ATCCGTTGCGGAGGAGGATGACCTGTGCGCTCGCGG 619  
Db 1655 CAACTTCGCAAGCGGGCATTCGCCCGCGGTACTTCG 1618

## RESULT 12

US-09-799-451-881  
; Sequence 881, Application US/09799451  
; Patent No. 6783969  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aiqiong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yundong

; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Ghosh, Reena  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/799,451  
; CURRENT FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 948  
; SOFTWARE: pc\_FL\_genes Version 2.0  
; SEQ ID NO 881  
; LENGTH: 4530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (330)..(4265)  
US-09-799-451-881

Query Match 4.3%; Score 57.6; DB 4; Length 4530;  
Best Local Similarity 46.7%; Pred. No. 0.00074;  
Matches 183; Conservative 0; Mismatches 209; Indels 0; Gaps 0;  
QY 140 TTTTCTACAAGCGCTCCAGGCGCGCGCGCGCATCCGAGACACCTTCTTCCCGCGTTCC 259  
Db 1384 TCTTCGTGTGGGACCGAGGCGAGAGGCTTGTGCGGAAGGCAAGGTCTCGGTGGAGTGG 1443  
QY 200 GGTTCCTCGCCACCGCGCGCTCCCGCGCGCATCCGAGACACCTTCTTCCCGCGTTCC 259  
Db 1444 TGACCGCTGGTGAGCGCCCTCCCGCGCGCTCTGCTCCCGCGCGCATCGGAGATCG 1503  
QY 260 ACGGCACCGACTCTCTCCCGAGGCGCGCGCGCGCATCTCTCACCTCTGCTCC 319  
Db 1504 AGGAAGGGAGATCGTCAGCGGAGGAGGCCAGGCTGGCGTGTCTCTTCTCCGCC 1563  
QY 320 TCGACGACCTCTCGCGGGTTTTCAGGCGCGCTCGCTCGCAGACATCAAGATCGCGCGCA 379  
Db 1564 CCGGCGCGCGCGCGCGCTTACACCGCGCGCTCGCGCACCCACCGCGCGCGCGCG 1623  
QY 380 TCAGTGGCCACCGAGTTTCGCGGAGCGCTCATCGCAAGTGCCTCGCAAGGACCGCG 439  
Db 1624 TTCTTCAGCTTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1683  
QY 440 GGACACGAGCGTTCTGTCTCGGATTCGCGCTTCGCGCGTTCGAGTCTCGCGCGCGCG 499  
Db 1684 AGTCGACGCGAGGCGCGCTTCGAGTGGACCTAGGGAGCGCGCTCCGCGCGCGCGCG 1743  
QY 500 CGCGCGTGTGCGGACGAGCGCGCGCGGAGGTG 531  
Db 1744 CCGCGACTCGCGTGGCG 1775

## RESULT 13

US-09-773-816-1  
; Sequence 1, Application US/09773816  
; Patent No. 6340774  
; GENERAL INFORMATION:  
; APPLICANT: Stanford University  
; APPLICANT: Khosla, Chaitan  
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR  
; FILE REFERENCE: 28600-20210.00  
; CURRENT APPLICATION NUMBER: US/09/773,816  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 60/243,458  
; PRIOR FILING DATE: 2000-10-25  
; PRIOR APPLICATION NUMBER: US 60/179,305  
; PRIOR FILING DATE: 2000-01-31



```

; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23673
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23623)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1

Query Match
Best Local Similarity 4.3%; Score 57.6; DB 3; Length 23673;
Matches 181; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 311 ACCTCGTCTCGACGACCTCTCTCGCGGGTTTCAGGCGCCCTCGTGCAGACATCAAGA 370
Db 14307 ACCCGGCGACTGCTGGCGGTGTTTCGCGCTGCGCGCTCTTTCGGATGCTGCCGATGC 14366

QY 371 TCGGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAAGTGCCTGCCA 430
Db 14367 GCGGCGACATCCGCGCCCAACCGGGGTGCTGAAGACCAACGACATGGTACGCTGGCG 14426

QY 431 AGGACCGGGGACACGAGGTTCTGCTCGATTTCGCGCTCTCCGGGTCCTCGAGTGTGCG 490
Db 14427 CGTCTCGCCGACGGCGGCGACCTTGACCTGCTGCTGACGTGCTGGCCCGCGCGCG 14486

QY 491 GCGCCGAGGCGCGTGTGCGGACGAGCGCCCGGAGGTGAAGGCCATGACACCCGCG 550
Db 14487 CCGACAAGGCGCGCGGTGCGCGCTTGACCTGCGCGCGCGCGCGGATGCTGGCG 14546

QY 551 GCGTCGCGCGCTGCTCGCGCGCTACGTGTCA---TCCGTTGCGGAGGAGGATGACT 607
Db 14547 GGTGCGCGTGGGCGTGTGCTGCGACGACCCGCTGTGCCGCTCGACGCGGAGTGGCG 14606

QY 608 GTGCGTCCGCGCGCGGTGTACGAGGAAAGTGGAGTCTGTGTCACAGCTCGCGGAGC 667
Db 14607 ACGTGTGCGCGCGCGGTGCGACGCCCTCAAGGCGACCGCGCTGCGACGCTGGCGAGC 14666

QY 668 TCAAGGCG 675
Db 14667 TCAGCGCG 14674

RESULT 14
US-09-105-537-32
; Sequence 32, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 11220
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-32

Query Match
Best Local Similarity 4.2%; Score 57; DB 3; Length 11220;
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 66 CCGCGCGGAGCACCAAGTCCGCGGCCACCGCGCTTCGCGGAGCGCGCGTGGGAGCACAGGT 185
Db 24009 CCGCGCGGAGCGCTCTGTGGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 24068

QY 126 CGACGGCTCGGCTCTTCTACAAGCGCTTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 24069 CGAGGCGCTGACGGGTGCTGCTGCTGATGCGGCGCTGCTGCTGCTGCTGCTGCTGCT 24128

QY 186 CGCTTCTATGAGCGCTTCTCCGCCCGCGCGCTTCGCGGAGCGCGCGCGCGCGCGCGCT 245
Db 24129 CGATGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24188

QY 246 CTTTCCCGCGGTTCACGCGACGCGACTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCG 305
Db 8322 CCGCGCGGACCGTCTGTGTGGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8381

```

```

126 CGACGGCTCGGCTCTTCTACAAGCGCTTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 185
8382 CGAGGCGCTGACGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8441

186 CGCTTCTATGAGCGGTTCGCGGCCACCGCGCGCTTCGCGGCGCGCGCGCGCGCGCGCGCTT 245
8442 CGATGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8501

246 CTTTCCCGCGGTTCACGCGACGCGACTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCA 305
8502 GCGGTTCACGCGGACGCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCGCGCGCGCGCG 8561

306 TCTACACCTCGCTCTCGACGACCTCTCGCGGGGTTTCAGGCGCGCGCTTCGCGCGCGCGCAT 365
8562 GGGCGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8621

366 CAAGATCGCGCGCATCAGTGGCCACCGAGTTCGCGGAGCGCGCGCGCGCGCGCGCGCGCT 425
8622 CCGGGACGCGCTGACGACCGGGGACCGCACCGTTCGCGCGCGCGCGCGCGCGCGCGCGCT 8681

426 CGCAAGGACCGCGGAGCGCGCGCTTCGCTGCGGTTTCGCGGTTTCGCGGTTTCGCGGTTTCG 485
8682 CGGACGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8741

486 CGTTCGCGCGCGCGCGCGCGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
8742 GGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8801

546 CGCGCGCGCTTCGCGCGCG 562
8802 CGCGCGCGCGCGCGCGCTCG 8818

RESULT 15
US-09-105-537-5
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match
Best Local Similarity 4.2%; Score 57; DB 3; Length 36778;
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 66 CCGCGCGGAGCACCAAGTCCGCGGCCACCGCGCTTCGCGGAGCGCGCGCGCGCGCGCGCG 125
Db 24009 CCGCGCGGAGCGCTCTGTGGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 24068

QY 126 CGACGGCTCGGCTCTTCTACAAGCGCTTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 24069 CGAGGCGCTGACGGGTGCTGCTGCTGATGCGGCGCTGCTGCTGCTGCTGCTGCTGCT 24128

QY 186 CGCTTCTATGAGCGCTTCTCCGCCCGCGCGCTTCGCGGAGCGCGCGCGCGCGCGCGCT 245
Db 24129 CGATGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24188

QY 246 CTTTCCCGCGGTTCACGCGACGCGACTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCG 305

```

Db	24189	GC	GTCC	ACGG	ACAG	CGCC	CGCT	CTGG	GGCC	CTCG	GGCC	GTCC	GC	GAC	GGAG	AGCC	24248	
QY	306	TC	CTAC	CTCG	TCTC	GAC	GA	CTCT	CG	GG	GTTC	AG	GG	CCCT	TC	GTCC	AGACAT	365
Db	24249	GG	CCCG	GTTC	GTCT	CTG	ACCT	CG	CC	GG	GAAG	CCCG	GA	CG	CG	GG	AGCC	24308
QY	366	CA	AGAT	CG	GG	CCAT	CAC	GTG	GC	CA	CCG	AGTT	CG	CC	GG	AGCC	TACAT	425
Db	24309	CG	GG	AG	CG	CC	TG	AC	CG	GG	AG	CG	CC	AC	CG	TC	GG	24368
QY	426	CG	CC	AA	GG	AC	CG	CG	GA	CC	AC	AG	CG	TT	CTG	TC	GA	485
Db	24369	CG	GC	AG	CG	CC	CT	CG	CA	CG	CC	CT	CG	GC	TC	CG	GC	24428
QY	486	CG	TC	GG	CC	CG	AG	GG	CG	CG	TC	GT	GG	CG	AG	CG	CG	545
Db	24429	GG	CG	TC	CT	CG	TAC	CC	CG	CG	CT	GG	CG	GG	CG	CG	CG	24488
QY	546	CG	CC	GG	CG	TC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	562
Db	24489	CG	GG	CC	GG	CG	CT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	24505

Search completed: June 17, 2005, 01:57:11  
Job time : 267 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 21:47:21 ; Search time 786 Seconds  
(without alignments)  
10122.306 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgagtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001s:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003s:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1344	6 AAD43514	Aad43514 Maize ino
2	864.4	64.3	922	6 AAD43513	Aad43513 Maize ino
3	844.8	62.9	1169	6 AAD43511	Aad43511 Maize ino
4	837.4	62.3	923	6 AAD43512	Aad43512 Maize ino
5	740.4	55.1	3416	6 AAD43522	Aad43522 Maize ino
6	686	51.0	899	6 AAD43518	Aad43518 Maize ino
7	470.4	35.0	643	6 AAD43519	Aad43519 Maize ino
8	298.4	22.2	519	6 AAD43520	Aad43520 Maize ino
9	191.6	14.3	353	6 AAD43521	Aad43521 Maize ino
10	184.6	13.7	1195	6 AAD43516	Aad43516 Eucalyptu
11	179	13.3	1105	6 AAD43515	Aad43515 Soybean i
12	149.8	11.1	1020	6 AAD43517	Aad43517 P. argent
13	148.6	11.1	1168	3 AAC39023	Aac39023 Arabidops
14	147	10.9	1104	3 AAC33685	Aac33685 Arabidops
15	147	10.9	1243	3 AAC38692	Aac38692 Arabidops
16	145.6	10.8	1130	3 AAC48750	Aac48750 Arabidops
17	142	10.6	1130	3 AAC40313	Aac40313 Arabidops
18	104.6	7.8	464	6 ABL93575	Ab193575 Arabidops
19	72.4	5.4	113193	8 AAD54645	Aad54645 Streptomy
20	71.4	5.3	135638	10 ABX34289	Abx34289 S. atrool

21	68.6	5.1	1383	10 ADG33751	Adg33751 Actinomyc
22	68.6	5.1	88421	6 AAL40781	Aal40781 8842int 9
23	65.4	4.9	484	10 ADE82058	Ade82058 Arabidops
24	65	4.8	897	8 ADA71279	Ada71279 Rice gene
c 25	63.8	4.7	5452	10 ADC86736	Adc86736 Human GPC
26	63.4	4.7	3849	4 AAF25795	Aaf25795 S. chrysos
27	63.4	4.7	9975	8 AAL61173	Aal61173 Actinosyn
c 28	63.4	4.7	82746	8 AAL61224	Aal61224 Actinosyn
29	63.2	4.7	1227	6 ABL61294	Ab161294 N. unifor
30	63.2	4.7	24379	2 AAT93095	Aat93095 Streptomy
31	63.2	4.7	24379	2 AAT93095	Aat93095 Streptomy
32	62.6	4.7	11619	13 ADQ91715	Adq91715 Polyketid
33	62.6	4.7	164051	13 ADQ91695	Adq91695 Polyketid
34	62.2	4.6	1084	10 ADB78935	Adb78935 Rice tran
35	62	4.6	2000	8 ADA71938	Ada71938 Rice gene
36	62	4.6	5877	6 ABS78681	Abs78681 Kitasatos
37	62	4.6	27541	4 AAD17185	Aad17185 Streptomy
38	62	4.6	125401	4 AAD17186	Aad17186 Streptomy
39	61.8	4.6	7788	10 ADI23894	Adi23894 Streptomy
40	61.8	4.6	37360	10 ADI23892	Adi23892 Streptomy
41	61.2	4.6	2475	12 ADJ44598	Adj44598 Plant cDN
42	61.2	4.6	3354	8 AAL61181	Aal61181 Actinosyn
43	61	4.5	2466	13 ADT43880	Adt43880 Bacterial
44	61	4.5	53789	2 AAV21187	Aav21187 Amycolato
45	61	4.5	64492	12 ADK16023	Adk16023 Streptomy

ALIGNMENTS

RESULT 1  
AAD43514  
ID AAD43514 standard; DNA; 1344 BP.  
XX  
AC AAD43514;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 52..921  
FT /\*tag= a  
FT /product= "Maize IPPK protein #4"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
XX 09-JAN-2002; 2002WO-US003120.  
XX  
XX 12-JAN-2001; 2001US-026146SP.  
XX  
XX (PTON-) PIONEER HI-BRED INT INC.  
XX  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX WPI: 2002-636540/68.  
XX P-PSDB; AAE26196.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
XX useful in modulating phytic acid biosynthesis by decreasing phytate or  
XX increasing non-phytate phosphorous to improve the nutritional value of  
XX animal feed.  
XX  
XX Claim 1; Page 64-65; 86pp; English.  
XX  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
XX

CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX

Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other;

Query Match 100.0%; Score 1344; DB 6; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 2.3e-271;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCACGAGTCAAGTCCGTCACCCCTCGCGCCCATAGTCCCTCCCATACATCCGAC 60  
DB 1 GCACGAGTCAAGTCCGTCACCCCTCGCGCCCATAGTCCCTCCCATACATCCGAC 60  
QY 61 CTCACCGCCGAGCACCAGTCCGCGCCACCGCGCTCCGCGAGCAAGCTGGGCGG 120  
DB 61 CTCACCGCCGAGCACCAGTCCGCGCCACCGCGCTCCGCGAGCAAGCTGGGCGG 120  
QY 121 CTCATCGACGGCTCGGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGTGGGGAGC 180  
DB 121 CTCATCGACGGCTCGGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGTGGGGAGC 180  
QY 181 GAGTCGCGCTTCTATGAGCGCTTCTCCGCCACGCGCGCTCCGCGCGCATCCGAGAC 240  
DB 181 GAGTCGCGCTTCTATGAGCGCTTCTCCGCCACGCGCGCTCCGCGCGCATCCGAGAC 240  
QY 241 ACCCTTCTCCCGCGTTCACGACGACGCTCTCCGCGGGTTTCAGGGCGCCCTGGTCGCA 300  
DB 241 ACCCTTCTCCCGCGTTCACGACGACGCTCTCCGCGGGTTTCAGGGCGCGACGCCCGGGAG 300  
QY 301 CGCATCTCACTCGTCTCGACGACCTCTCGCGGGTTTCAGGGCGCCCTGGTCGCA 360  
DB 301 CGCATCTCACTCGTCTCGACGACCTCTCGCGGGTTTCAGGGCGCCCTGGTCGCA 360  
QY 361 GACATCAAGATCGGCGCATCACTGGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAG 420  
DB 361 GACATCAAGATCGGCGCATCACTGGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAG 420  
QY 421 TGCCTCGCAGGACCGGACACACGAGCTTCTCGGATTCGCGCTCCGCGCTC 480  
DB 421 TGCCTCGCAGGACCGGACACACGAGCTTCTCGGATTCGCGCTCCGCGCTC 480  
QY 481 CGAGTCGTCCGCGCGCGCGCTGTGGCGGACGAGCGCGGAGGTGAAGCCATG 540  
DB 481 CGAGTCGTCCGCGCGCGCGCTGTGGCGGACGAGCGCGGAGGTGAAGCCATG 540  
QY 541 GACACCGCGCGCTCCGCGCGCTGTCCGCGCTACGTTGTCATCCGTTGCCGAGAGGG 600  
DB 541 GACACCGCGCGCTCCGCGCGCTGTCCGCGCTACGTTGTCATCCGTTGCCGAGAGGG 600  
QY 601 ATGAGCTGTGCGCTCCGCGCGCGTGTACGAGGAGAAAGGTGAGTCTTGTCAAGCTG 660  
DB 601 ATGAGCTGTGCGCTCCGCGCGCGTGTACGAGGAGAAAGGTGAGTCTTGTCAAGCTG 660  
QY 661 CGCAGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 720  
DB 661 CGCAGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 720  
QY 721 CTTCTGGGCTATGATGCTGCTGACGTCGAGCGGAGATGGGGGTGGGGTGACCGTG 780  
DB 721 CTTCTGGGCTATGATGCTGCTGACGTCGAGCGGAGATGGGGGTGGGGTGACCGTG 780  
QY 781 AAGCTGGTGGACTTGGCCATGTGCGAGGGTGAATGGGATGACCAACTTCTG 840  
DB 781 AAGCTGGTGGACTTGGCCATGTGCGAGGGTGAATGGGATGACCAACTTCTG 840  
QY 841 GCGCGGCTCTGCTCGCTCATCAAGTTCGTTCTGACATTTGTCGAGACTCTTCATACG 900  
DB 841 GCGCGGCTCTGCTCGCTCATCAAGTTCGTTCTGACATTTGTCGAGACTCTTCATACG 900

DB 841 GCGCGGCTCTGCTCGCTCATCAAGTTCGTTCTGACATTTGTCGAGACTCTTCATACG 900  
QY 901 CAGCCTTGGGTCCTTCTTAAGAGAGATCCTCGCATTTTCGATTGATAAAGCCCTA 960  
DB 901 CAGCCTTGGGTCCTTCTTAAGAGAGATCCTCGCATTTTCGATTGATAAAGCCCTA 960  
QY 961 CAAGTTTGTCTGAAAAAGAGCGCTCCGAGTTGTGCTGGGTGTGAGATCTGAGACG 1020  
DB 961 CAAGTTTGTCTGAAAAAGAGCGCTCCGAGTTGTGCTGGGTGTGAGATCTGAGACG 1020  
QY 1021 GTCGTGCGCCCATCTGTTGCTTGGCTTGGCTTGGCTGCAACATACGGAACCTGCT 1080  
DB 1021 GTCGTGCGCCCATCTGTTGCTTGGCTTGGCTTGGCTGCAACATACGGAACCTGCT 1080  
QY 1081 CCTTTTTCGCAACCCCTTACTTCCGAGAAACTTTTTCCTCCACTTTGGGGGTTCGA 1140  
DB 1081 CCTTTTTCGCAACCCCTTACTTCCGAGAAACTTTTTCCTCCACTTTGGGGGTTCGA 1140  
QY 1141 TTACGTTGGATCTGGTTTGGCCACTCGCTCAGAGTTGTAAAGCATGAGGAGGCGTGT 1200  
DB 1141 TTACGTTGGATCTGGTTTGGCCACTCGCTCAGAGTTGTAAAGCATGAGGAGGCGTGT 1200  
QY 1201 TGATCCGCGCAACTGTGTAGTCTTTGCGCTGCGCTGCGCTTTCGATGGCTTTTGCCTGC 1260  
DB 1201 TGATCCGCGCAACTGTGTAGTCTTTGCGCTGCGCTGCGCTTTCGATGGCTTTTGCCTGC 1260  
QY 1261 TCGATCCGATGTGTACTGAGATCGTAGTGATGAGCGTCTTACCTCCAAACGAATCCG 1320  
DB 1261 TCGATCCGATGTGTACTGAGATCGTAGTGATGAGCGTCTTACCTCCAAACGAATCCG 1320  
QY 1321 TCCGATAAAAAA 1344  
DB 1321 TCCGATAAAAAA 1344  
RESULT 2  
AAD43513  
ID AAD43513 standard; DNA; 922 BP.  
XX AC AAD43513;  
XX DT 14-NOV-2002 (first entry)  
XX DE Maize inositol polyphosphate kinase (IPPK) DNA #3.  
XX KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
XX OS Zea mays. nutritional value; animal feed; transgenic; gene; ds.  
FH Key Location/Qualifiers  
CDS 53..922  
FT /\*tag= a  
FT /product= "Maize IPPK protein #3"  
XX  
PN WO200259324-A2.  
XX  
XX 01-AUG-2002.  
XX 09-JAN-2002; 2002WO-US003120.  
XX 12-JAN-2001; 2001US-0261465P.  
XX (PTON-) PIONEER HI-BRED INT INC.  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX WPI; 2002-636540/68.  
XX P-FSDB; AAE26195.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
XX useful in modulating phytic acid biosynthesis by decreasing phytate or  
XX increasing non-phytate phosphorous to improve the nutritional value of



```
QY 162 CGGACCGTGGGAGCAGAGTGGCTTCTATGAGGCGTTCCTCGGCCACGCGCGGT 221
DB 194 CGGACCGTGGGAGCAGAGTGGCTTCTATGAGGCGTTCCTCGGCCACGCGCGGT 253
QY 222 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTCCACGGGACGCGACTCTCTCCCCAC 281
DB 254 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTCCACGGGACGCGACTCTCTCCCCAC 313
QY 282 CGAGCGCAGCCCGGGAGCGCATCTCTCACTCGTCTCTCGACGACTCTCTCGGGGGTT 341
DB 314 CGAGCGCAGCCCGGGAGCGCATCTCTCACTCGTCTCTCGACGACTCTCTCGGGGGTT 373
QY 342 TCAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCCATCACTGCGCCACCGAGTTGCGC 401
DB 374 TGAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCCATCACTGCGCCACCGAGTTGCGC 433
QY 402 GGAGCCCTACATCGCAAGTGGCTCGCCAAAGACCGCGGGACCAAGCGTTCCTGCTCGG 461
DB 434 GGAGCCCTACATCGCAAGTGGCTCGCCAAAGACCGCGGGACCAAGCGTTCCTGCTCGG 493
QY 462 ATTCGCGTCTCGCGGTCGAGTGGTGGGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 521
DB 494 ATTCGCGTCT--TGCGTCCGAGTGGTGGGCGCGGCGGCGGCGGCGGCGGCGGCGG 551
QY 522 CCGGAGGTGAAGGCATGGACACCGCGCGGTCGCGCGGTCGCTCGGCGCTACGTGTC 581
DB 552 CCGGAGGTGAAGGCATGGACACCGTCGCGGTCGCGCGGTCGCTCGGCGCTACGTGTC 611
QY 582 ATCCG-TTGGCGACGAGGGATGGACTGTGGCTCGCGCGGCGGCTGTACGGAGGAAAG 640
DB 612 ATCCGTTGGCGACGAGGGATGGACTGTGGCTCGCGCGGCGGCTGTACGGAGGAAAG 671
QY 641 GTGAGTCTTGTACAGTGGCGAGCTCAAGGCGTGGTTCGAGGAGCAGACTCTGTTC 700
DB 672 GTGAGTCTTGTACAGTGGCGAGCTCAAGGCGATGGTTGGAGGAGCAGACTCTGTTC 731
QY 701 ACTTCTACTCGCGTCGATCTTCTGGGCTATGATGCTGCTGCGAGTCGACGCGCGGAG 760
DB 732 ACTTCTACTCGCGTCGATCTTCTGGGCTATGATGCTGCTGCGAGTCGACGCGCGGAG 791
QY 761 ATGGGGTGGGTGACGCTGAAGCTGGTGGACTTTGGCCATGTGGCCGAGGCTGATGGG 820
DB 792 GTGGGGTGGGTGACGCTGAAGCTGGTGGACTTTGGCCATGTGGCCGAGGCTGATGGG 851
QY 821 TGATTGACCAACTTCTCGCGGGCTCTGCTGCTGATCAAGTTCTGATGATG 880
DB 852 TGATTGACCAACTTCTCGCGGAGCTCTGCTAGCTGATCAAGTTCTGATGATG 911
QY 881 TTCCGGAGACTCTCATAGCGAGCCTTTGGGTCCTTCTTAAGAGGATCTTGGA-TTT 939
DB 912 TTCCAGAGACTCTTATAGCGAGCCTTTGGGTCCTTCTTAAGAGGATCTTGGAATTT 971
QY 940 CGATTGTGATAACAAAG 955
DB 972 TGATTGTATACAAAG 987
```

## RESULT 4

```
AAD43512
ID AAD43512 standard; DNA; 923 BP.
XX
AC AAD43512;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #2.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Zea mays.
XX
```

```
PH Key Location/Qualifiers
FT CDS 53..736
FT /*tag= a
FT /product= "Maize IPPK protein #2"
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
XX WPI; 2002-636540/68.
XX
XX P-PSDB; AAE26194.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 1; Page 60-61; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
XX Sequence 923 BP; 138 A; 326 C; 286 G; 173 T; 0 U; 0 Other;
```

```
Query Match 62.3%; Score 837.4; DB 6; Length 923;
Best Local Similarity 97.5%; Pred. No. 1.9e-165;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY 42 TCCCATACCATGTCGCGACCTCCACCGCGGAGCACAAGTCCGCGCCACCGCGCTC 101
DB 43 TCCCATACCATGTCGCGACCTCCACCGCGGAGCACAAGTCCGCGCTCACCAGCGCTC 102
QY 102 CGCCAGCAAGTGGGCGCGCTCATCGACGGCTCCGGGCTCTTCTACAAGCGCTCCAGGC 161
DB 103 CGCAGCAGCGCGGCGCGCTCATCGAGGGCTCCGGGCTCTTCTACAAGCGCTCCAGGC 162
QY 162 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGGCGTTCCTCGGCCACGCGCGGT 221
DB 163 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGGCGTTCCTCGGCCACGCGCGGT 222
QY 222 CCGGCGCGCATCCGAGACACTTCTCCCGGTTCCACGGGACGCGACTCTCTCCCCAC 281
DB 223 CCGGCGCGCATCCGAGACACTTCTCCCGGTTCCACGGGACGCGACTCTCTCCCCAC 282
QY 282 CGAGCGCAGCGCGGAGCGCATCTCTCATCTCGTCTCGACGACTCTCTCGGGGGTT 341
DB 283 CGAGCGCAGCGCGGAGCGCATCCGACCTCGTCTCGACGACTCTCTCGGGGATT 342
QY 342 TCAGGCGCGCTGCTCGCAGACATCAAGATCGGCGCATCACTGTTGGGCGACCGAGTTGCC 401
DB 343 TGAGGCGCGCTGCTCGCAGACATCAAGATCGGCGCATCACTGTTGGGCGACCGAGTTGCC 402
QY 402 GGAGCGCTACATCGCAAGTGGCTCGCGCAAGGACCGGGGACCAAGCGTTCGTCTCGG 461
DB 403 GGAGCGCTACATCGCAAGTGGCTCGCGCATGACCGGGGACCAAGCGTTCGTCTCGG 462
QY 462 ATTCGCGGTCTTCCGCGCTCGAGTCTCGGCGCCCGGAGGCGCGCTGTGTGGCGAGCG 521
```

Db 463 ATTCCGCGTCTCCGCGTCCGAGTCCGTCCTCCCGGAGGCGCGGTGGCGGACGAGCG 522  
 Qy 522 CCCGAGGTGAAGGCATGACACCGCCGCGGTCCGCGGTGCTCCGCGCTACGTGTC 581  
 Db 523 CCCGAGGTGAAGGCATGACACCGCCGCGGTCCGCGGTGCTCCGCGCTACGTGTC 582  
 Qy 582 ATCCG-TTGGCCGACGAGGGATGGAAGTGGCTGGCGCTCCGCGCGGTGTACGAGGAAAG 640  
 Db 583 ATCCGCTTCCGACGAGGGATGGAAGTGGCTGGCGCTCCGCGCGGTGTACGAGGAAAG 642  
 Qy 641 GTGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGTGTTCCAGGAGCAGACTCTGTTC 700  
 Db 643 GTGAGTCTTGTCACTGTGCGGAGCTCAAGCGGTGTTCCAGGAGCAGACTCTGTTC 702  
 Qy 701 ACTTCTACTCGCGTCCGATCTTCTGGGCTATGCTGCTCCAGTCCGACGAGCGGAG 760  
 Db 703 ACTTCTACTCGCGTCCGATCTTCTGGGCTATGATGCTGCTGCAAGTCCGACGAGCGGAG 762  
 Qy 761 ATGGGGTGGGTGACGCTGAAGCTGGTGGACTTTGGCCCATGTGGCCGAGGTGATGGG 820  
 Db 763 GTGGGGTGGGTATACAGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGTGATGGG 822  
 Qy 821 TGATTGACCACAACCTTCTCGGCGGCTCTGCTCGCTGATCAAGTTCGTTCTGACATTG 880  
 Db 823 TGATTGACCACAACCTTCTCGGCGGCTCTGCTAGCTGATCAAGTTCGTTCTGACATTG 882  
 Qy 881 TTCGGAGACTCTCTATACGAGCGCTTTGGGTCTCTTTAA 921  
 Db 883 TTCCAGAGACTCTCTCAGACGAGCGCTTTGGGTCTCTTTAA 923

## RESULT 5

AAD43522  
 ID AAD43522 standard; DNA; 3416 BP.  
 XX  
 AC AAD43522;  
 XX  
 DT 14-NOV-2002 (first entry)  
 XX  
 DE Maize inositol polyphosphate kinase (IPPK) DNA #9.  
 XX  
 KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic; gene; ds.  
 XX  
 OS Zea mays.

Key Location/Qualifiers  
 FT 72..407  
 CDS /\*tag= a  
 FT /product= "Maize IPPK protein #6"  
 FT

WO200259324-A2.

PN 01-AUG-2002.  
 XX  
 PD 09-JAN-2002; 2002WO-US003120.  
 XX  
 PF 12-JAN-2001; 2001US-0261465P.  
 XX  
 PR (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PA Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 PI WPI; 2002-636540/68.  
 XX  
 DR P-PSDB; AAE26201.  
 XX

PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 XX  
 PS Claim 1; Page 75-77; 86pp; English.

CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK DNA

XX Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;

Query Match 55.1%; Score 740.4; DB 6; Length 3416;  
 Best Local Similarity 90.2%; Pred. No. 4.4e-145;  
 Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4;

Qy 2 CACGAGTCAAGTCCGCTCACCCCTCGCGCCATAGTCCCTTCCCACATACCATGTCGAC 61  
 Db 28 CACCATCGCCACCCGTCACCCCTTGTCTCCATAG-----TCCCACATACCATGTCGAC 81  
 Qy 62 TCCACCCGCGGAGCACCAAGTCCGCGCCACCGCTCCGCGCAGCAAGCTTGGSCCGC 121  
 Db 82 TCCACCCGCGGAGCACCAAGTCCGCGCTCACCGGCTCCGCGCAGCAAGCTTGGSCCGC 141  
 Qy 122 TCATCGACGCTCCGCGCTCTTTACAAGCGCTCCAGCGCGCGACCGTGGGAGCAG 181  
 Db 142 TCATCGACGCTCTGGGCTCTTTACAAGCGCTCCAGCGCGCGACCGTGGGAGCAG 201  
 Qy 182 AGTTCGCTTCTATGAGCGCTTCCGCGCACGCGCGCTCCGCGCCGCGATCCGAGACA 241  
 Db 202 AGTTCGCTTCTATGAGCGCTTCTCCGCCACGCGCGCTCCGCGCCGCGATCCGAGACA 261  
 Qy 242 CCTTCTCCCGCGTCCAGCGCACGCGACTCTCTCCACCGAGCGCAGCCCGGGAGC 301  
 Db 262 CCTTCTCCCGCGTCCAGCGCACGCGACTCTCTCCACCGAGCGCAGCCCGGGAGC 321  
 Qy 302 CGCATCTCTACCTCGTCTCGACGACCTCTCGCGGGGTTCAGCGCGCTTCGTCGAG 361  
 Db 322 CGCATCTCGACCTCTCTCGACGACCTCTCTCGCGGGGTTCAGCGCGCTTCGTCGAG 381  
 Qy 362 ACATCAAGATCGCGCCATCAAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCAGT 421  
 Db 382 ACATCAAGATCGGTGCCATCAAGT-----405  
 Qy 422 GCCTCGCCAGGACCGCGGAGCACGAGCGTTCGTCGGATTCCGCGCTCTCCGCGCTCC 481  
 Db 406 -----GACCACGAGCGTTCGTCGGATTCCGCGCTCTCCGCGCTCC 446  
 Qy 482 GAGTCTCGGCGCCGAGGCGCGCTGTGGCGGACGAGCGCGCGAGGTGAAGGCCATGG 541  
 Db 447 GAGTCTCGGCGCCGAGGCGCGCTGTGGCGGACGAGCGCGCGAGGTGAAGGCTATGG 506  
 Qy 542 ACACCGCGCGTCCGCGCGTGTCCGCGCTACGTCGATTCGCG- TTGCGGACGAGGG 600  
 Db 507 ACATTGTCCGCGTCCGCGCGTGTCCGCGCTACGTCGATTCGCGCTTTCGCGACGAGGG 566  
 Qy 601 ATGAGCTGTGCTCCGCGCGTGTACGAGGAAAGCTTGTGTACAGCTG 660  
 Db 567 ATGAGCTGTGCTCCGCGCGTGTACGAGGAAAGCTTGTGTACAGCTG 626  
 Qy 661 CGCAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 720  
 Db 627 CGCAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 686  
 Qy 721 CTTCTGGCTATGATGCTGTCAGTCCACGACGCGGAGATGGGGTGGGGTGAACGCTG 780  
 Db 687 CTTCTGGCTATGATGCTGTCAGTCCACGACGCGGAGGTGGGGTGAACGCTG 746  
 Qy 781 AAGCTGTGGACTTTGCCCATGTCGCGAGGGTGTATGGGGTGTATGACCAACATCTCTG 840  
 Db 747 AAGCTGTGGACTTTGCCCATGTCGCGAGGGTGTATGGGGTGTATGACCAACATCTCTG 806  
 Qy 841 GCGGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTGTTCCGAGAGACTCCTCATACG 900

```
Db 807 GCGGGGCTCTGCTAGCTGATCAAGTTGTTCTGACATTTGTTCCAGAGATCTCTCAGCG 866
Qy 901 CAGCCTTTGGGTCTTCTTAAGAGAGGATCTGCA-TTTCGATTTGATAACAAAG 955
Db 867 CAGCCTTTGGGTCTTCTTAAGAGAGGATCTGCAATTTTCGATTTGATAACAAAG 922

RESULT 6
AAD43518
ID AAD43518 standard; DNA; 899 BP.
XX
AC AAD43518;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #5.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 89..424
FT /*tag= a
FT /product= "Maize IPPK protein #4"
XX
PN WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX PR 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
XX WPI; 2002-636540/68.
XX
XX P-PSDB; AAE26200.
XX
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 1; Page 73-74; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
XX Sequence 899 BP; 140 A; 312 C; 276 G; 171 T; 0 U; 0 Other;
XX
XX Query Match 51.0%; Score 686; DB 6; Length 899;
XX Best Local Similarity 88.7%; Pred. No. 8.4e-134;
XX Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;
XX
Qy 2 CACGAGGTCACTCCGTACACCCCTCGGCGCCATAGTCCCTTCCCATACCATGTCGGACC 61
Db 45 CACATCGCCACCCGTACCCCTTGTCTCCATAG-----TCCCATACCATGTCGGACC 98
Qy 62 TCCACCCGCGGAGCACCAGTCTCGCGGCCACCGCGCTCCGCCAGCAAGTGGGCGCGC 121
Db 99 TCCACCCGCGGAGCACCAGTCTCGCGGCCCTCACCAGTCTCGGCCAGCAAGTGGGCGCAC 158
```

```
Qy 122 TCATCAGCGCTCGGCTCTTCTTAAGCCGCTCCAGCGCGGACCGCTGGGAGCAGC 181
Db 159 TCATCAGCGCTCTGGCTCTTCTTAAGCCGCTCCAGCGCGGACCGCTGGGAGCAGC 218
Qy 182 AGGTCCGCTTCTATGAGGGGTTCTCCGCCACACCGCCGCTCCCGGCCGATCCGAGACA 241
Db 219 AGGTCCGCTTCTATGAGGGGTTCTCCGCCACACCGCCGCTCCCGGCCGATCCGAGACA 278
Qy 242 CTTTCTTCCCCCGTTTCCAGGCAACGCACTCTCTCCCAACCGAGGCGGAGCCCGGGAGC 301
Db 279 CTTTCTTCCCCCGTTTCCAGGCAACGCACTCTCTCCCAACCGAGGCGGAGCCCGGGAGC 338
Qy 302 CGCATCTCACCTCGTCTCGACGACCTCTCTCGGGGGTTTTCAGGGCCCTCTCGTCCGAG 361
Db 339 CGCATCTCGTACCTCGTCTCGACGACCTCTCTCGGGGGTTTTCAGGGCCCTCTCGTCCGAG 398
Qy 362 ACATCAAGATCGGCGGCATCACGTGGCCACCGAGTTTCGCGGAGCCCTACATCGCCAAGT 421
Db 399 ACATCAAGATCGTGCATCACGT-----
Qy 422 GCCTCGCCAAGSACCGCGGACCAACGAGCGTTCTGTCTCGGATTCGCGTCTCGGGCTCC 481
Db 423 -----GACCATGAGCGATCTGCTCGGATTCACAGTCTCGGGCTCC 463
Qy 482 GAGTCTCGGCCCCGAGGGGCGCTGTGGCGGACGAGCGCCCGGAGGTGAAGGCCATGG 541
Db 464 GAGTCTGTGGCCCCGAGGGGCGCTGTGGCGGACGAGCGCCCTGAGGTGAAGGCCATGG 523
Qy 542 ACACCGCGGCGTCCGCGCGTCTCCGGCGTACGTTGTCTCATCCG-TTGCAGCAGAGGG 600
Db 524 ACATTGTGGCGTCCGCGCGTCTCGGGCGTGCATGTCTCATCCGTTGCCGCGAGGG 583
Qy 601 ATGGAATGTGCGTCTCGCGCGGTGTACGAGGAAAGGTGGAAGTCTTTGTCAACGCTG 660
Db 584 ATGGAATGTGCGTCTCGCGCGGTGTACGAGGAAAGGTGGAAGTCTTTGTCAACGCTG 643
Qy 661 CGGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGGGCTCGATT 720
Db 644 CGGAGCTCAAGCGGTGTTTCGAGGAGGCGAGACTCTGTTCCACTTCTACTCGGGCTCGATT 703
Qy 721 CTTCTGGGCTATGATGCTGTCAGTTCGAGGCGGAGATGGGGTGGGGTGACGGTG 780
Db 704 CTTCTGGGCTATGATGCTGTCAGTTCGAGGCGGAGTGGGGTGGGGTAACAGTG 763
Qy 781 AAGCTGTGGAATTTTGCCTCATGTGCGCGGAGGTGATGGGGTGAATTGACCAACTTCTG 840
Db 764 AAGCTGTGGAATTTTGCCTCATGTGCGCGGAGGTGATGGGGTGAATTGACCAACTTCTG 823
Qy 841 GCGGGGCTCTGCTCGCTGATCAAGTTGTTTCTGACATTTCTCCGAGACTCTCTCATACG 900
Db 824 GCGGGGCTCTGCTAGTGTGATCAAGTTGTTTCTGACATTTCTCCAGAGACTCTCTTAGACG 883
Qy 901 CAGC 904
Db 884 CAGC 887

RESULT 7
AAD43519
ID AAD43519 standard; DNA; 643 BP.
XX
AC AAD43519;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #6.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; ds.
XX
OS Zea mays.
XX
```



PN WO200259324-A2.  
XX 01-AUG-2002.  
XX 09-JAN-2002; 2002WO-US003120.  
XX PF  
XX PR 12-JAN-2001; 2001US-0261465P.  
XX PA (PION-) PIONEER HI-BRED INT INC.  
XX PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX DR WPI; 2002-636540/68.  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
XX Claim 1; Page 74; 86pp; English.  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
XX Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;  
SQ  
Query Match 35.0%; Score 470.4; DB 6; Length 643;  
Best Local Similarity 95.9%; Pred. No. 9.5e-89;  
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;  
Qy 42 TCCCATACCATGTCGACATCCACCGCGGAGCACCAAGTCCCGCCACCGGCTC 101  
Db 107 TCCCATACCATGTCGACATCCACCGCGGAGCACCAAGTCCCGCTACCGGCTC 166  
Qy 102 CGCCAGCAAGTGGGCGCGCTCATCGAGCGTCCGGCTCTTACAGCGCGTCCAGGC 161  
Db 167 CGCCAGCAAGTGGGCGCGCTCATCGAGCGTCCGGCTCTTACAGCGCGTCCAGGC 226  
Qy 162 CGGCGACCGTGGGAGACAGAGTCCGCTTCTATGAGCGGTTCTCGGCCACCGCGCGT 221  
Db 227 CGGCGACCGTGGGAGACAGAGTCCGCTTCTATGAGCGGTTCTCGGCCACCGCGCGT 286  
Qy 222 CCGGCGCGCATCCGAGACATCTTCCCGCGGTTCCACGCGCGGACATCTCCCGCAC 281  
Db 287 CCGGCGCGCATCCGAGACATCTTCCCGCGGTTCCACGCGCGGACATCTCCCGCAC 346  
Qy 282 CGAGCGCAGCCCGGAGCGCGCATCTCTACCTCGTCTCGACGACCTCTCTCGGGGTT 341  
Db 347 CGAGCGCAGCCCGGAGCGCGCATCTCTCGTCTCTCGACGACCTCTCTCGGGGTT 406  
Qy 342 TCAGGCGCCCTGCGTGGCAGATCAAGATCGGCGCCATCACTGCGGCCACCGAGTTGCC 401  
Db 407 TGAGGCGCCCTGCGTGGCAGATCAAGATCGGCGCCATCACTGCGGCCACCGAGTTGCC 466  
Qy 402 GGAGCCCTACATCCCAATGCTCTCGCAAGACCGGCGGACCGAGCGTTCTGCTCGG 461  
Db 467 GGAGCCCTACATCCCAATGCTCTGCGCAAGACCGGCGGACCGAGCGTTCTGCTCGG 526  
Qy 462 ATTCCGCTCTCCGCGCTCCGAGTCCGCGCCCGCGGCGCGGTCGCGGACGAGCG 521  
Db 527 ATTCCGCTCT--TGCGTCCGAGTCTGCGGCCCGCGGCGCGGTCGCGGACGAGCG 584  
Qy 522 -CCGCGAGTGAAGSCCATGGACA--CGGCCGCGGCTCCGCGCGGTCGCTCCGG 570  
Db 585 CCGCGGCGGTGAANGCTATGACACACCCGTCGCGNGNCGCGCGNGTCTGCTCGG 635

RESULT 8  
AAD43520  
ID AAD43520 standard; DNA; 519 BP.  
XX  
XX AAD43520;  
AC  
XX 14-NOV-2002 (first entry)  
DT  
XX  
XX Maize inositol polyphosphate kinase (IPPK) DNA #7.  
DE  
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; de.  
KW  
XX  
XX Zea mays.  
OS  
XX  
XX WO200259324-A2.  
PN  
XX  
XX 01-AUG-2002.  
PD  
XX  
XX 09-JAN-2002; 2002WO-US003120.  
PF  
XX  
XX 12-JAN-2001; 2001US-0261465P.  
PR  
XX  
XX (PION-) PIONEER HI-BRED INT INC.  
PA  
XX  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
PI WPI; 2002-636540/68.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
XX Claim 1; Page 75; 86pp; English.  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
XX Sequence 519 BP; 144 A; 109 C; 126 G; 125 T; 0 U; 15 Other;  
SQ  
Query Match 22.2%; Score 298.4; DB 6; Length 519;  
Best Local Similarity 89.4%; Pred. No. 8.5e-53;  
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;  
Qy 627 GTACGGAGGAAAGTGGAGTCTTGTCAAGTCCGAGCTCAAGCGTGGTTCGAGGA 686  
Db 2 GTACGGAGGAAAGTGGAGTCTTGTCAAGTCCGAGCTCAAGCGTGGTTCGAGGG 61  
Qy 687 GCAGACTCTGTTCCACTTCTACTCGGCGTCAATCTTCTGGGCTATGATGCTGTCAGT 746  
Db 62 GCAGACTCTGTTCCACTTCTACTCGGCGTCAATCTTCTGGGCTATGATGCTGTCAGT 121  
Qy 747 CGCAGCAGGCGGAGTGGGCTGGGCTGACCGTGAAGCTGGTGAATTTGCCCATGTGGC 806  
Db 122 CGCAGCAGGCGGAGTGGGCTGGGCTGACCGTGAAGCTGGTGAATTTGCCCATGTGGC 181  
Qy 807 CGAGGGTGATGGGTGATGACCAAACTTCTCTGGGCGGGCTCTGCTGCTGATCAAGTT 866  
Db 182 CGAGGGTGATGGGTGATGACCAAACTTCTCTGGGCGGGCTCTGCTGCTGATCAAGTT 241  
Qy 867 CGTTTCTGACATTTGTTCCGAGACTCTCTACACGAGCTTTGGGTCCTTCTTAAGAGAG 926  
Db 242 TGTTTCTGACATTTGTTCCGAGACTCTCTACACGAGCTTTGGGTCCTTCTTAAGAGAG 301

```
QY 927 GATCCTGGCA-TTTCGATTGATAACCAAGCCCTACAAAGTTTCTCTCGGAAAGAGCG 985
Db |||||
QY 302 GATCCTGGCATTTTCGATTGATAACCAAGGAANCATTTTCAGTCGCCAAAAAANCA 361
Db |||||

QY 986 CCTCCGA 992
Db |||||
362 CCAAGTA 368

RESULT 9
AAD43521
ID AAD43521 standard; DNA; 353 BP.
XX
AC AAD43521;
XX
XX 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #8.
XX
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; ds.
XX
XX Zea mays.
XX
XX WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 75; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
SQ Sequence 353 BP; 81 A; 78 C; 91 G; 96 T; 0 U; 7 Other;
Query Match 14.3%; Score 191.6; DB 6; Length 353;
Best Local Similarity 90.2%; Pred. No. 1.7e-30;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCAGCGGTGTCGAGGAGCAGACTCTGTTCACTTCTACTCGGGGTCGATTTCTG 726
Db |||||
1 CTCAGGCGATGGTTGGAGGAGCAGACTCTGTTCCACTTCTACTCGGGGTCGATTTCTG 60

QY 727 GGCTATGATGCTGTCAGTCGACGAGCGGAGATGGGGTGAGCGTGAAGCTG 786
Db |||||
61 GGCTATGATGCTGTCAGTCGACGAGCGGAGGATGGGGTGAGCGTGAAGCTG 120

QY 787 GTGACTTTGCCCATGTGGCCGAGGGGTGATGGGT--GATTGACCAAACTTCTGGGGG 844
Db |||||
121 GTGGACTTTGCCCATGTGGCCGAGGGGTGATGGGTGATTTTGTGACCAAACTTCTGGGGG 180
```

```
QY 845 GGCTCTGCTCGCTGATCAAGTTC--GTTTCTGACATTGTTCCGAGACTCCTCATACG-C 901
Db |||||
181 AGCTCTGCTAGCTGATCAAGTTCGTTTCTTGACATTGTTCCAGANACTCCTTAGAGGCC 240

QY 902 AGCCTTTGGTGCCTCTCTTAAGAGA 925
Db |||||
241 AGCCTTTGGTGCCTCTCTTAAGAAA 264

RESULT 10
AAD43516
ID AAD43516 standard; DNA; 1195 BP.
XX
AC AAD43516;
XX
XX 14-NOV-2002 (first entry)
XX
DE Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.
XX
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
KW nutritional value; animal feed; gene; ds.
XX
XX Eucalyptus grandis.
XX
XX Key Location/Qualifiers
XX CDS 116..1148
XX FT /*tag= a
XX FT /product= "E. grandis IPPK protein"
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAB26198.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 68-70; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is Eucalyptus grandis IPPK DNA
XX
SQ Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;
Query Match 13.7%; Score 184.6; DB 6; Length 1195;
Best Local Similarity 55.6%; Pred. No. 6e-29;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY 56 CCGACCTCCACCGCGGAGACCAAGTCGCGGCCACCGCCTCGCCAGCAAGCTGG 115
Db |||||
114 CCATGCTCAAGGTCCCGGATCATCAAGTCGCGGTCAACGGGAGACGGGGGAAGCTGG 173

QY 116 GCCCGCTCATGACGCGCTCGGCTCTTCTACAAGCGGCTCCAGGCGCGACCGTGGGG 175
Db |||||
```

Db 174 GGCCACTGGTGGATGATTGGGCGCTCTTATAGCCTCTCCAGAGCGATCATCGCGAG 233  
Qy 176 AGCAGAGTGGCTTCTATGAGCGTTCTCCGCCACGCGCGTCCCGCCGATCC 235  
Db 234 ACACGGAAGTGGCTTTACAGTCAATTTATCCATACGAGATCCAGGTCACATTC 293  
Qy 236 GAGACACTTCTTCCCGGTTCCACGCGACGAGTCTCTCCGCCAGCGCGAGCCG 295  
Db 294 G---CAAAATCTTCTCGGTTTACCGAATAAGACTAT---TGAGCGCTCTGATGAT 347  
Qy 296 GGGAGCGCATCTCACTCGTCTCCAGCACTCTCTCGCGGGTTTACGCGCCCTGCG 355  
Db 348 CGGTCTCTCAACCTCACTCTGTTCTGAGGATCTCTGTCGGGTGCGACCAACCAATCTC 407  
Qy 356 TCGCAGATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCCGCGAGCCCTACATCG 415  
Db 408 TCATGGACATCAAGACTGGATCCAGAACATGGTATCCGGAGGCTCTGAGGAGTACATCC 467  
Qy 416 CCAAGTGCCTCGCCAAAGACCGCGGACCAAGCGTCTCTCGGATTCGCGCTCTCCG 475  
Db 468 AAAAGTCTTAGAAGAGATCGAATAGCACAGCGTTTCATTGGGTTTGGATTTCTG 527  
Qy 476 GCGTCCAGTGTCTCG---CCCCGAGGCGCGGTGTGGCGACGAGCGCGCGAGGTGA 532  
Db 528 GGCTAAGGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAGAGTTCTTT 587  
Qy 533 AGGCATGGACACCGCGCGGTCCCGCGGTCTCCCGCGTACGTCTCATCC----- 585  
Db 588 ATAGCTTTAATGCGGACGGTGTACAGTCTGAGGAAAGTTTGTCTTCCAACTTGT 647  
Qy 586 ----GTTGCGACGAGGGATGACTGTGCGTCTCGCGCGGTGTACGAGGAAAG 640  
Db 648 CTCTGGGTCAAAATGTGATCCGATGTTGTTGATGATCAAAAGTTTACTGTCAACCGG 707  
Qy 641 GTGAGTCTTTGTACAGCTGCGGAGCTCAAGCGGTGTTGAGGAGCAGACTCTGTTC 700  
Db 708 GTGGAATTTTGGCACAATTTGCTCAGCTGAAGGAATGGTTGAGGTTTCAAGCAATATC 767  
Qy 701 ACTTCTACTGGGTCAATTTCTTGGCTATGATGCTGTGAGTGCAGCAGCGCGAG 760  
Db 768 ACTTCTATTCTGTCTCACTCATTTATCTATATGACAGGAGTCTGTGACGCG--- 823  
Qy 761 ATGGGGTGGGTGACGCTGAAGCTGTGGAATTTTGCCATGTGCGCGAGGGTCAATGGG 820  
Db 824 --TGTGCACACCCGAAAGTTAACTGTGTGACTTTGACATGTGATGATGCGCGCG 881  
Qy 821 TGATTGACCAAACTTCTGCGCGGGTCTGCTCGCTGATCAAGTTGTTTCTGACATTG 880  
Db 882 TGATCGATCACAACTTCTTGGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941  
Qy 881 TTCCGGA 887  
Db 942 CTGATGA 948

RESULT 11

AAD43515  
ID AAD43515 standard; DNA; 1105 BP.  
XX AC  
XX AC AAD43515;  
XX AC  
DT 14-NOV-2002 (first entry)  
XX Soybean inositol polyphosphate kinase (IPPK) DNA.  
DE DE  
XX Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
XX nutritional value; animal feed; transgenic; gene; ds.  
XX Glycine max.  
XX OS  
XX Key Location/Qualifiers  
XX FH 12..851  
XX CDS /\*tag= a  
FT

/product= "Soybean IPPK protein"  
W0200259324-A2.  
01-AUG-2002.  
09-JAN-2002; 2002WO-US003120.  
12-JAN-2001; 2001US-0261465P.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX WPI; 2002-636540/68.  
XX P-PSDB; AAE26197.  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
XX useful in modulating phytic acid biosynthesis by decreasing phytate or  
XX increasing non-phytate phosphorous to improve the nutritional value of  
XX animal feed.  
XX Claim 1; Page 66-67; 86pp; English.  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
XX polypeptides and polynucleotides. Sequences of the invention are useful  
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
XX increasing non-phytate phosphorous to improve the nutritional value of  
XX animal feed, or to reduce the environmental impact of animal waste.  
XX Polynucleotides of the invention are to produce transgenic plants with an  
XX altered phenotype. IPPK proteins are used to screen compounds that  
XX modulate their activity and raising anti-idiotypic antibodies. The  
XX present sequence is soybean IPPK DNA  
XX SQ Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;  
Query Match 13.3%; Score 179; DB 6; Length 1105;  
Best Local Similarity 55.6%; Pred. No. 8.8e-28;  
Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;  
Qy 61 CTCACCGCGCGGAGACCAAGTCCGCGCACCGGCTCCGCGCAGAGCTGGGCGG 120  
Db 15 CTCGAAGATCCCGGAGCACAGGTGGCGGCGCACAAAGCCAGGAACTCTGGGCCA 74  
Qy 121 CTCATGACGGCTCCGCGCTCTTCTACAAGCGCTCCAGCGCGCGCGCGCGGAGCAC 180  
Db 75 CTCGTGACGATTTTGGAAAATTTCTACAAGCCCCCTCCAGACCAACAAGACGACGACC 134  
Qy 181 GAGTGGCTCTTATGAGGCGTT-----CTCGCGCCACCGCGCTCCGCGCGCGCTACTCC 234  
Db 135 CGCGGCTCCACCGNACTCTCTTTTACACCTCTCTCGCGCGCGCGCGCGCTACTCC 194  
Qy 235 CGAGACACTTCTTCCCGCGGTTCCAGCGACGCGACTCTCTCCCGCGCGCGCGCGCG 294  
Db 195 ATCGGTCTCTTCTTCCCGCGCTTTACGCGACCGCGCTCTCTGAGCGCTCTCGAGCGTCC 254  
Qy 295 GGGGAGCGCGATCTCTACCTCTCTCTGAGCGCTCTCTCGCGGGTTTCAGGCGCCCTGC 354  
Db 255 GG---TCCCGACCTCTACCTCTCTCTGAGGAGCTCTCTCTGCGGCTACTCTCAAAACCTCC 311  
Qy 355 GTGCGAGACATCAAGATCGCGCGCATCACTGGCGCACCGAGTTTCGCGGAGCCCTACATC 414  
Db 312 GTCATGACGTAAGATCGGCTCCAGAACCTTGGACCTGGGAGACTCCGAGGACTACATC 371  
Qy 415 GCCAAGTCTTCGCCAAGGACCGCGGACCAAGCGGCTTCTGTCGAGATTCCGCGTCTCC 474  
Db 372 TGCAAGTCTTGAAGAAGGACAGAGTCTCTAGTCTTGGCTTGGGTTTCAGAAATCTCG 431  
Qy 475 GGGTCCGAGTCTGTCGCGCGCGGCGCGCTGTGGCGGAGCGCGCGCGCGGAGGTGAAG 534  
Db 432 GGAGTCAAGGACTCTATCTCTCTCTGGAACCTTACCAGGAAATCTCTCAGTGTCTATCC 491  
Qy 535 GCCATGACACCGCGCGCGTCTCGCGCGCTCTCGCGGCT---ACGTGTCTATCGTTCGCC 591

```
Db 492 GCCCATGGTGTGCACCTTGTCTCAACAAGTTCTGTTCTCTTAATATATACACCATGAT 551
QY 592 GACGAGGGATGAGTGTGCGCTCGCCGCGCGGTGTACGGAGGAAAAAGGTGGAGTCTTG 651
Db 552 GATCATCATCCGATTGCGCTTTTCGCAACGAGGTCTAC-----GGCGCGGTTTG 602
QY 652 TCACAGCTGGCGAGCTCAAGGCGTGGTTCGAGGACAGACTGTGTCACCTTCTACTCG 711
Db 603 GAGCGGTTCGAGAAGCTCAAGGACTGGTTGCGAGGTTTCAGACGGTGTATCATTCTATTCT 662
QY 712 GCCTCGATTCTTCTGGGCTATGATGCTGTGCAGTTCGACGCGGCGAGATGGGGTGG 771
Db 663 TGTTCGTCTTGTGTGTACGAGAAGGATCTAGGAAA-----GGGAAAGTACCAAC 716
QY 772 GTGACGGTGAAGTGTGTGACTTTGCCCCATGTCGCCGAGGGTGTACGGGTGATGTACAC 831
Db 717 CCTCTGGTCAAACTCGTTGACTTTGCACACGTTGGTGGACGGAACGGTGTCAATTGATCAC 776
QY 832 AACTTCTGGCGGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACAT 878
Db 777 AACTTCTGGGTGGCTTTGTTCTTCATCAAGTTCTCAAGGATAT 823
```

## RESULT 12

AAD43517 standard; DNA; 1020 BP.

```
XX AAD43517;
AC AAD43517;
DT 14-NOV-2002 (first entry)
DE P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.
KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
KW nutritional value; animal feed; gene; ds.
XX Parthenium argentatum.
XX Key Location/Qualifiers
FH 21..908
CDS /*tag= a
FT /product= "P. argentatum IPPK protein #1"
FT
FT
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26199.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 1; Page 71-72; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
```

```
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is P. argentatum IPPK DNA
XX
SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;
Query Match 11.1%; Score 149.8; DB 6; Length 1020;
Best Local Similarity 53.2%; Pred. No. 1.1e-21;
Matches 456; Conservative 0; Mismatches 362; Indels 39; Gaps 5;
QY 61 CTCACCCCGCCGGAGCACCAAGTCGCGGCCACCGCGCTCCAGCGCGGACCGTGGGAGCAC 180
Db 24 CTCAAGGCCCCAGATCATCAGGTTCGTGACATGAAGCTGGGCTCGGAAGCTTGGCCCCA 83
QY 121 CTCATCGAGCGGTCCGGCTCTTCTACAAGCGCTCCAGCGCGGACCGTGGGAGCAC 180
Db 84 CTCATGTGATTCAGGCGCGTTTACAAACACATGCGGGGTGATTAACCGTGGGTGAGAA 143
QY 181 GAGGTGCGCTTCTATGAGCGGTTCTCCGCCACCGCGCTCCCGCCCGCATCCGAGAC 240
Db 144 GAAGTAGCCTTTTATGAATCATTTTCTTACATAATATATTCAGAAACACATACG--C 200
QY 241 ACCTTCTTCCCGGTTTCCACGCGACGCACTCTCTCCCGACGAGCGCGAGCCCGGGAG 300
Db 201 AAATTCCTTCTATATATATGCGCACCAAAAT-----CATGAAGGCATCCACTGGCTCT 254
QY 301 CCGCATCTCTCAGCTCGTCTCGACGACCTCTCGGGGGTTCAGCGCGCTCGCTCGCA 360
Db 255 GACCATCTCTCATATGTTGCAAGATCTTATCAGCTCATGTCAACCATCTGTAATG 314
QY 361 GACATCAAGATCGGGCCATCACTGGCCACCGAGTTCCCGGAGCCCTACATCGCAAG 420
Db 315 GACATCAAAATCGGTTCCAGAACATGCGCGCCAGAACTTCCGAGCGGTACATTGCAAA 374
QY 421 TGCCTCGCCAAAGACCGCGGACCAAGCGTTCGTCTCGATTCGCGTCTCCCGCGTC 480
Db 375 TGCCTAAAAAAGGATAGGAAAGCACAAGTATTCATTCGGATTCAGGATCTCCCGGCTG 434
QY 481 CGAGTCGTGCGCCCGGAGCGCGCTGTGGCGGAGG---AGCGCCCGGAGGTGAAGGCC 537
Db 435 CAAGTCTATATCATGATGATGGGTCAAGGTTTATAGCCCTCATAGAAATTACATGCGTAA 494
QY 538 ATGGACACCGCGGGTCCGCGCGTCTCGCGCGTCTAGCTCATCTCGTTGCGGACGAG 597
Db 495 ACCGCGCCAGCTGATGTTAGACTACTTCTTAGGAAATTTGTTTCTTAACCGTCTGCA 554
QY 598 GGGAT-----GGACTGTGCGCTCGCCCGCGGCGGTGTTACGCA 633
Db 555 GAGATGGAAATGCGCACAGGCCCTAGGCCCGGATTTGTTCTTTAGCATCTTTTGTATGGT 614
QY 634 GGAAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGTGGTTTCGAGGAGCAGACT 693
Db 615 GGGCTAATGGGATATTAGCTCAACTGATGGAATTGAAGACATGGTTTGAAGATCAACA 674
QY 694 CTGTTCCACTTCTACTCGCGCTCGATTTCTTCTGGGCTATGATGCTGCTGCAGTCGACGA 753
Db 675 ATTTACCACTTCTATGCTTGTCTTTTGTGTCATCTTTGAAAGAGGTTGGTTGTTAAA 734
QY 754 GCGGAGATGGGGGTGGGTGACCGTGAAGCTGGTGGACTTTTCCCATATGTCGCGAGGTT 813
Db 735 GGT--GCTCGGTTCAAAACGACAGAAAGTCAAACTTATTTGATTTTGTCTCATGTATACAG 791
QY 814 GATGGGTGATTTGACCACAACTTCTTGGCGGGCTGCTGCTCGCTGATCAAGTTCTGTTCT 873
Db 792 AATGGTGTATTGATCACAATTTCTTTGGGTGGGCTCTGTCTTTGTAAGATTCAATTTCT 851
QY 874 GACATTGTTCCGGAGAC 890
Db 852 GACATCTTTCGGAGAC 868
RESULT 13
AAC39023
ID AAC39023 standard; DNA; 1168 BP.
```

XX AAC39023; PR 18-JUN-1999; 99US-0139462P.  
AC 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
DT 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
KW 23-JUN-1999; 99US-0140354P.  
KW 24-JUN-1999; 99US-0140695P.  
KW 28-JUN-1999; 99US-0140823P.  
XX 29-JUN-1999; 99US-0140991P.  
OS 30-JUN-1999; 99US-0141287P.  
XX 01-JUL-1999; 99US-0141842P.  
XX 02-JUL-1999; 99US-0142154P.  
XX 06-JUL-1999; 99US-0142055P.  
XX 08-JUL-1999; 99US-0142390P.  
XX 08-JUL-1999; 99US-0142803P.  
XX 09-JUL-1999; 99US-0142920P.  
XX 12-JUL-1999; 99US-0142977P.  
XX 13-JUL-1999; 99US-0143542P.  
XX 14-JUL-1999; 99US-0143624P.  
XX 15-JUL-1999; 99US-0144005P.  
XX 16-JUL-1999; 99US-0144085P.  
XX 16-JUL-1999; 99US-0144086P.  
XX 19-JUL-1999; 99US-0144325P.  
XX 19-JUL-1999; 99US-0144331P.  
XX 19-JUL-1999; 99US-0144332P.  
XX 19-JUL-1999; 99US-0144333P.  
XX 19-JUL-1999; 99US-0144334P.  
XX 19-JUL-1999; 99US-0144335P.  
XX 20-JUL-1999; 99US-0144352P.  
XX 20-JUL-1999; 99US-0144352P.  
XX 21-JUL-1999; 99US-0144884P.  
XX 21-JUL-1999; 99US-0144814P.  
XX 21-JUL-1999; 99US-0145086P.  
XX 21-JUL-1999; 99US-0145088P.  
XX 22-JUL-1999; 99US-0145085P.  
XX 22-JUL-1999; 99US-0145087P.  
XX 22-JUL-1999; 99US-0145089P.  
XX 22-JUL-1999; 99US-0145192P.  
XX 23-JUL-1999; 99US-0145145P.  
XX 23-JUL-1999; 99US-0145218P.  
XX 23-JUL-1999; 99US-0145224P.  
XX 26-JUL-1999; 99US-0145276P.  
XX 27-JUL-1999; 99US-0145913P.  
XX 27-JUL-1999; 99US-0145918P.  
XX 27-JUL-1999; 99US-0145919P.  
XX 27-JUL-1999; 99US-0145951P.  
XX 02-AUG-1999; 99US-0146386P.  
XX 02-AUG-1999; 99US-0146388P.  
XX 02-AUG-1999; 99US-0146389P.  
XX 03-AUG-1999; 99US-0147038P.  
XX 04-AUG-1999; 99US-0147204P.  
XX 04-AUG-1999; 99US-0147302P.  
XX 05-AUG-1999; 99US-0147192P.  
XX 05-AUG-1999; 99US-0147260P.  
XX 06-AUG-1999; 99US-0147303P.  
XX 06-AUG-1999; 99US-014716P.  
XX 08-AUG-1999; 99US-0147493P.  
XX 09-AUG-1999; 99US-0147835P.  
XX 10-AUG-1999; 99US-0148171P.  
XX 11-AUG-1999; 99US-0148319P.  
XX 12-AUG-1999; 99US-0148341P.  
XX 13-AUG-1999; 99US-0148565P.  
XX 13-AUG-1999; 99US-0148684P.  
XX 16-AUG-1999; 99US-0149368P.  
XX 17-AUG-1999; 99US-0149175P.  
XX 18-AUG-1999; 99US-0149426P.  
XX 20-AUG-1999; 99US-0149722P.  
XX 20-AUG-1999; 99US-0149723P.  
XX 20-AUG-1999; 99US-0149929P.  
XX 23-AUG-1999; 99US-0149902P.  
XX 23-AUG-1999; 99US-0149930P.  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23077.  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EPI033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 05-MAR-1999; 99US-0123588P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139452P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.

PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 11.1%; Score 148.6; DB 3; Length 1168;		
Best Local Similarity 53.0%; Pred. No. 2e-21;		
Matches 453; Conservative 0; Mismatches 369; Indels 33; Gaps 5;		
QY	49	ACCATTCGCACTCCACCGCGGAGCACCAAGTCGCGGCGCACCGCGCTCCGCGAGC 108
Db	170	AACAAGATGCAGCTCAAGATCCCTGAACATCAGGTTGCAGGACACATTGCTAAAGACGGG 229
QY	109	AGCTGGGCGGCTCATCCAGCGCTCCGCGCTTCTATAGCGCTTCTCCGCCCCACCGCGCGTCCGCGGAC 168
Db	230	AGCCTGGTCTCTCGTAGATCAAGGGTCGGTCTTCAAGCCACTTCAGGCGGATTC 289
QY	169	CGTGGGAGCAGGAGTCGCTTCTATAGCGCTTCTCCGCCCCACCGCGCGTCCGCGGCG 228
Db	290	CGTGGTGAATCGAGGTAAAGTCTACGAATCTTCTCTCAACACAGAGGTTCCAGAA 349
QY	229	CGATCCGAGACACCTTCTTCCCGGTTCCAGCGACGCGACTCTCCCGACCGAGGCG 288
Db	350	CACATCC---ATAGATATTTCGCGGTGTATCAGCGCACTCAA-----GCAGTT 394

QY	289	CAGCCCGGGAGCGCATCCTCCTCGTCCGACGACCTCCTCGCGGGTTTCAGGCG 348
Db	395	GAAGGTTCTGATGGAGCAGCCATGATGGTGTGGAAATCTTCTTGCAATACTCAAAA 454
QY	349	CCCTGCGTCGACAGATCAAGATCGGCGGCATCAGGTGCGCCACCGAGTTTCGCGGAGCC 408
Db	455	CCATCAGTAATGATGTTAAGATGGTTCACAGAACATGATCCTGATGCATCGAAGAA 514
QY	409	TACATCGCAAGTCCCTCGCCAAAGACCGCGGACCAAGAGGTTCTGCTCGGATTCGCG 468
Db	515	TACATCCAAAAATGTTTGAAGAAAGACACGGGTACCAACAACGGTGTCTATCGGGTTTCAGG 574
QY	469	GTCTCGGCGTCCGAGTGTCTCGCGCCCGAGGGCGC---GTGTGGCGGACGAGCCCGCG 525
Db	575	ATCTCTGGTTTCAAGTGTATGATCACAAGAAATCGAGTTTCTGGAAGCCCGAGAGAAAG 634
QY	526	GAGGTGAAGGCCATGAGACACACCGCGCGTCCCGCGTCTCCGCGCTACGTTGTC--- 582
Db	635	CTTCTTCGCGGGCTCGATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 694
QY	583	-----TCCGTTGCGGACGAGGGATGGAATGTGCGCTCGCGCGGGTGTACGGA 633
Db	695	AACTCAGCTTTCGGACACTGGCTCGAAACCTGACTCTGCTTTTGCCTCGAGTGTTCACGGC 754
QY	634	GGAAAAAGGTGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGTGGTTCGAGGAGCAGACT 693
Db	755	GTTTCCGACGGATCTTAACGCGATTGTCGAACTCAAGACTGGTTCGAGAACCAACG 814
QY	694	CTGTTCCACTTCTACTCGCGCTCGATCTTCTTGGGCTATGATGCTGCTGCAGTC---GCA 750
Db	815	CTCTACCATTTCAACTCTTGTTCGATTTTAATGGTCTATGAGAATGAATCCATCTTGAAG 874
QY	751	GCAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAG 810
Db	875	GGAAATGATGATGATGCTAGACCAACAGTCAAGCTGGTGGATTTTCTCATGTTCTTGAT 934
QY	811	GCTGATGGGTGATTGACCACAACTTCTTGGCGGGCTCTGCTCGCTGATCAAGTTTCGTT 870
Db	935	GGTAATGGTGCATTGACCACTAATCTTTGGGTGGTCTTTGCTCTTTTATAAATTCATT 994
QY	871	TCTGACATTGTTCCG 885
Db	995	CGTGAGATTCTTCAG 1009
RESULT 14		
AAC33685		
ID	AAC33685 standard; DNA; 1104 BP.	
XX	AAC33685;	
AC	AAC33685;	
XX	17-OCT-2000 (first entry)	
DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 3949.	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 3949.	
XX	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic pathway;	
KW	promoter; termination sequence; ss.	
XX	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
PN	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-00301439.	
PF	25-FEB-1999; 99US-0121825P.	
XX	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	

PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0132456P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140355P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.

PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	22-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 10.9%; Score 147; DB 3; Length 1104;			
Best Local Similarity 52.9%; Pred. No. 4.3e-21;			
Matches 452; Conservative 0; Mismatches 370; Indels 33; Gaps 5;			
Qy	49	ACCATGTCGAGACTCCACCGCGGAGCACCAAGTCCGCCGACCGCCCTCCGCGCAGC 108	
Db	120	AACAAGATGCAGCTCAAGATCCCTGAACATCAGGTTGCAGGACACATTGCTAAAGACGGG 179	
Qy	109	AGCTGGGCGGCTCATCGAGCGCTCCGGCTCTTTCTAGAGCGGCTCCAGCGGCTCCAGCGCGGAC 168	
Db	180	AGCCCTGGTCTCTCGTAGATGACAAGGGTGGGTTCTTCAAGCCACTTCAGGCGGATTTCT 239	
Qy	169	CGTGGGGAGCACGAGGTCGGCTTCTATGAGCGGTTCTCCGCCACCGCCCGCTCCCGGCC 228	
Db	240	CGTGGTGAATCGAGGTAAGTTCTACGAACTTTTCTCTCAACACAGAGGTTCCAGAA 299	
Qy	229	CGCATCCGAGACACTTCTTCCCGCGGTTCCAGCGCACGCGACTCTCTCCCGCCAGCGGC 288	
Db	300	CACATCC---ATAGATATTCCCGGTGTATCAGCGCACTCAA-----GCAGTT 344	
Qy	289	CAGCCCGGGAGCGGCATCTCACTCGTCTCGACGACTCTCTCGCGGGGTTTCAGGCG 348	
Db	345	GAAGGTTCTGATGGAGCAGCCATGATGGTGTGGAAATCTTCTTCAGAAATACTCAAAA 404	
Qy	349	CCCTCGGTCGACACATCAAGATCGCGGCATCAGCTGGCGCACCGAGTTTCGCGGAGGCC 408	
Db	405	CCATCAGTAATGATGTTAAGATGGGTTTGAGNACATGTTATCTGATGCATCTGAAGAA 464	
Qy	409	TACATCCCAAGTCCCTCGCCAAAGGACCGGGACCAAGAGCGTTCTGCTCGGATTTCCGC 468	
Db	465	TACATCCAAAAATGTTTGAAGAAGACACGGGTACCACAAACCGTGTATCGCGGTTTCAGG 524	
Qy	469	GTCTCCGGGTCCGAGTCGTGCGCCCGCGAGGGGCC---GTGTGGCGGACGAGCGCCCG 525	
Db	525	ATCTCTGGTTTTCGAAGTGTATGATCAACAAGATTCGAGTTTCTGGAAGCCCGCAGAGGAAG 584	
Qy	526	GAGGTGAAGCCATGACACCGCGCGGCTCCCGCGGCTCGCGGGCTACGTTGTCA--- 582	
Db	585	CTTCTTCGGCGGCTCGATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 644	
Qy	583	-----TCCGTTCCGACGAGGGATGGAAGTGTGCGCTCGCGCGCGGCGGTGTAAGGA 633	
Db	645	AACTCACTTTTCGGACACTGGCTCGAAAACTGACTCTGCTTTTTCGCTCGAGTGTTTACGGC 704	

Qy	634	GGAAAAGGTGGAGTCTTGTCTACAGCTCGCGAGCTCAAGGCGTGGTTCGAGGAGCAGACT 693	
Db	705	GGTTCCACGGGATCTTAACGCGATGCTGGAACTCAAGACCTGGTTCGAGAACCAACG 764	
Qy	694	CTGTTCCACTTCTACTCGGCGTFCGATTCTTCTGGGCTATGATGCTGCTGCAGTC---GCA 750	
Db	765	CTCTACCATTTCAACTCTTGTTCGATTTTAATGGTCTATGAGAATGAATCCATCTTGAAG 824	
Qy	751	GCAGCGGAGATGGGGGTGGGTGACCGGTGAAGCTGGTGGACTTTTGCCCATGTGGCCGAG 810	
Db	825	GGAAATGATGATGATGCTAGACCACAAAGTCAAGCTGGTGGATTTTGCTCATGTTCTTGAT 884	
Qy	811	GGTGATGGGTGATTGACCAACAACCTTCTGGCGGGCTCTGCTCGCTGATCAAGTTCGTT 870	
Db	885	GGTAATGGTGTCAATTGACCAATAAATCTTGGGTGGTCTTTGCTCTTTCAAAACCTTCATT 944	
Qy	871	TCTGACATTGTTCCG 885	
Db	945	CGTGAGATCTTCAG 959	
RESULT 15			
AAC38692			
ID	AAC38692 standard; DNA; 1243 BP.		
XX	XX		
AC	AAC38692;		
XX	XX		
DT	17-OCT-2000 (first entry)		
XX	XX		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 21894.		
XX	XX		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	XX		
OS	Arabidopsis thaliana.		
XX	XX		
PN	EP1033405-A2.		
XX	XX		
PD	06-SEP-2000.		
XX	XX		
PF	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		
PR	18-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		



PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 28-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.

PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 10.9%; Score 147; DB 3; Length 1243;		
Best Local Similarity 52.9%; Pred. No. 4.4e-21;		
Matches 452; Conservative 0; Mismatches 370; Indels 33; Gaps 5;		
QY	49	ACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGC 108
DB	257	AACAAGATGAGCTCAAGTCCCTGAACATCAGGTTGCAGGACACATTGCTAAAGACGGG 316
QY	109	AGCTGGGCGGCTCATCGAGCGCTCCGCGCTTCTTACAAGCCGCTCAGGCCGGCGAC 168
DB	317	AGCCTGGTCTCTCGTAGACAAGGGTCGGTCTTCAAGCCACTTCAGGGCGATTCT 376
QY	169	CGTGGGAGCAGAGGTGCGCTTCTATAGGGGTCTTCCGCCACCGCCCGCTCCCGGC 228
DB	377	CGTGTGAATCGAGGTAAGTCTACGAATCTTCTCTCAACACAGAGGTTCCAGAA 436
QY	229	CGCATCCGAGACACTTCTTCCCGGGTTCACGGCACCGACTCCTCCCCACCGAGCG 288
DB	437	CACATCC---ATAGATATTCCCGGTGATACGGCACTCAA-----GCAGTT 481
QY	289	CAGCCGGGAGCGGATCCTCACCTCGTCTCGAGGACTCTCTCGGGGGTTTCAGGG 348
DB	482	GAAGGTTCTGATGGAGCAGCCATGATGGTGTGGAAATCTTCTTGCAGATACTCAAAA 541
QY	349	CCCTGGTCGCGAGACATCAAGATCGGCGCATCACGTGGCCACCGAGTTCGCGGAGCC 408
DB	542	CAATCAGTAATGGATGTTAAGATGGTTCGAGAACATGGTATCCTGATGCATCTGAAGAA 601
QY	409	TACATCGCAAGTGCTCGCCAAAGACCGCGGACCCAGAGGTTCTGCTCGGATTCGCG 468
DB	602	TACATCCAAAATGTTTGAAGAAAGACACGGGTACCCACACCGTGTCATCGGGTTTCAGG 661
QY	469	GTCTCGGCGTCGAGTCTCGGCCCGGAGGGGGCC---GTGTGGGAGCGAGGCGCCG 525
DB	662	ATCTCTGGTTCGAAGTGATGATCAAAAGAATCGAGTTTCTGGAAGCCCGAGAGGAAG 721
QY	526	GAGGTGAAGGCCATGACACCGCGCGCTCCGCGGTCTCCGGCTACGTGTCA--- 582
DB	722	CTTCTTCGCGGCTCGATGATAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 781
QY	583	-----TCCGTTGCCGACGAGGGGATGGAATGTGCGCTCGCCGCGCGGTGACGGA 633
DB	782	AACTCACTTTCGACACTGGCTCGAAACCTGACTCTGCTTTTGCCTCGAGTGTTTACGGC 841
QY	634	GGAAAGGTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGTGTTCGAGGAGCAGACT 693
DB	842	GGTCCACGGGATCTTAAACGAGTTGTGGAACCTAAGACCTGGTTCGAGAACCAACG 901
QY	694	CTGTTCCACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGCTGACGTC---GCA 750
DB	902	CTCTACCATTTCAACTCTTGTTCGATTTTAATGGTCTATGAGAAATCACTCTTGAAG 961
QY	751	GCAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGCCCGAG 810
DB	962	CGAAATGATGATGCTAGACCAAGTCAAGCTGGTGGATTTTGTCTCATGTCTTGTAT 1021
QY	811	GGTATGGGGTATTGACCAACTTCTTGGGGGGCTCTGCTCGCTGATCAAGTTCGTT 870
DB	1022	GGTAATGGTGTCTATTGACCAACTTCTTGGGGGTCTTTGTCTCTTTTCATAAACTTCATT 1081
QY	871	TCTGACATTGTTCCG 885
DB	1082	CGTGAGATTCTTCAG 1096